

当代西方经济学经典译丛

主编／周黎安

冲突的战略

The Strategy of Conflict

[美] 托马斯·谢林／著

《冲突的战略》

2005年诺贝尔经济学奖得主

托马斯·谢林开创性巨著！

華夏出版社

華夏出版社

VISIT…

LANZAROTE
Caliente.COM

2005 年诺贝尔经济学奖得主

托马斯·谢林开创性巨著!

当代西方 经济学 经典译丛

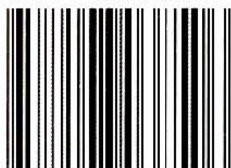
The Strategy of Conflict

冲突的战略

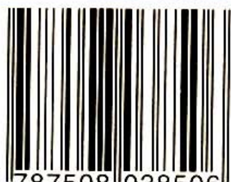
"为什么有些国家、团体和个人可以和平地解决冲突，而另一些国家、团体和个人却不断地被冲突困扰呢？感谢谢林和奥曼的研究，为这一自古以来困扰我们的问题带来启迪。"

诺贝尔经济学奖评委会主席　约尔根·韦布尔

"经济学大师的原创贡献大都体现为高度技术性的论文，让一般的读者望而却步。所幸谢林的原创贡献是一个例外，他的《冲突的战略》没有涉及多少专业术语和数学模型，完全依靠逻辑的力量展现思想的深度和广度，无疑是体现博弈论如何帮助人们理解现实世界的最



ISBN 7-5080-3850-9
定价: 32.00 元



ISBN 7-5080-3850-9
定价: 32.00 元

冲突的战略

The Strategy of Conflict

[美] 托马斯·谢林 / 著
赵 华 等 / 译
郑志刚 王 勇 / 审校

华夏出版社

图书在版编目(CIP)数据

冲突的战略/(美)谢林著;赵华等译.
-北京:华夏出版社,2006.1
ISBN 7-5080-3850-9

Ⅰ.冲… Ⅱ.①谢… ②赵… Ⅲ.对策论-应用-社会学 Ⅳ.C91

中国版本图书馆 CIP 数据核字(2005)第 127930 号

Thomas C. Schelling: *The Strategy of Conflict*
Chinese translation Copyright©2005 by HUAXIA PUBLISHING HOUSE
Original English language edition Copyright©1960,1980 by President and Fellows of Harvard College. C
本书英文版由 HARVARD UNIVERSITY PRESS 于 1980 年出版。
本书中文简体版权由 HARVARD UNIVERSITY PRESS 授予华夏出版社,版权为华夏出版社所有。未经出版者书面允许,不得以任何方式复制或抄袭本书内容。

版权所有,翻印必究。

北京市版权局著作权合同登记号:01-2004-2622

冲突的战略

〔美〕托马斯·谢林 著
赵华等 译

责任编辑: 李欣利
出版发行: 华夏出版社
(北京市东直门外香河园北里 4 号 邮编:100028)
经　　销: 新华书店
印　　刷: 世界知识印刷厂
版　　次: 2006 年 1 月北京第 1 版
2006 年 1 月北京第 1 次印刷
开　　本: 670×970 1/16 开
印　　张: 17.5
字　　数: 280 千字
插　　页: 2
定　　价: 32.00 元

本版图书凡印刷、装订错误,可及时向我社发行部调换

本 PDF 电子书制作者：

阿拉伯的海伦娜

爱问共享资料首页：

http://iask.sina.com.cn/u/1644200877

内有大量制作精美的电子书籍！！！

完全免费下载！

进入首页，点击“她的资料”，你就会进入一个令你惊叹的书的海洋！

当然，下载完了你理想的书籍以后，如果你能留言，那我将荣幸之至！



目　录

第一部分　战略理论构成要素

第 *1* 章　缓慢发展的国际战略学 …… *3*

第 *2* 章　论谈判 …… *19*

谈判实力：自我约束的实力 …… *20*
谈判机制及其结构特点 …… *25*
威　胁 …… *31*
允　诺 …… *37*
博弈游戏 …… *40*

第 *3* 章　谈判，沟通与有限战争 …… *47*

默契（利益共同）…… *48*
默式谈判（利益冲突）…… *52*
显式谈判 …… *59*
默式谈判和有限战争 …… *65*
前期准备 …… *67*

第二部分　博弈论的重新定位

第 *4* 章　趋向互动决策理论 …… *73*

博弈再分类 …… 77
协作博弈 …… 78
混合博弈中的暗示与互动认知 …… 86

第5章 实施、沟通与战略行为 …… 101

行 为 …… 103
威 胁 …… 104
允 诺 …… 111
主动权的放弃 …… 116
定 位 …… 117
授 权 …… 119
调 停 …… 120
沟通及其破坏性 …… 122
博弈矩阵中的行为组合 …… 125
矛盾的战略优势 …… 131
“战略行为” …… 132

第6章 博弈论与实验研究 …… 135

第三部分 战略：充满随机性

第7章 允诺和威慑的随机性 …… 145

失败的风险 …… 147
威胁意外实施的风险 …… 151
随机性承诺 …… 152

第8章 产生随机因素的威慑 …… 155

意外战争威胁 …… 156
有限战争：风险的根源 …… 158

有限战争中的危险行为 …… 160
报复与侵扰 …… 161
风险行为和“强制性”威胁 …… 162
边缘政策 …… 165
不完全决策过程 …… 167

第四部分　突袭：一种互不信任的分析

第 9 章　双方对突袭的担心 …… 173
概率的无限排序 …… 174
具有“明确答案”的非合作博弈 …… 175
重温旧题 …… 182
不完全预警系统产生的随机行为 …… 184
参变行为 …… 187
默式博弈 …… 189
谈判博弈 …… 190
多人博弈 …… 191
第 10 章　突袭与裁军 …… 193
误解侵袭 …… 206
有限战争中的误解 …… 207
互存误解 …… 207
长期监控 …… 209
超前机制 …… 209

附　录

附录 A　核武器与有限战争 …… 215

附录 *B* 博弈论的非对称性 …… *223*

附录 *C* 重释“非合作”博弈的决策理念 …… *243*

冲突利益 …… *250*
第三者的影响 …… *252*
释义收益 …… *253*
选手数量 …… *253*
结　论 …… *254*

英中文译名对照及索引 …… *255*

译后记 …… *263*

序 言

2005年瑞典皇家科学院授予托马斯·谢林诺贝尔经济学奖，以表彰他在博弈论方面的开创性贡献，尤其是在增进人们对合作和冲突问题的理解方面的重要工作。与谢林分享这一殊荣的还有以色列经济学家罗伯特·奥曼（Robert Aumann）。

托马斯·谢林1921年生于加州奥克兰，1951年获得哈佛大学经济学博士学位，先后在耶鲁大学、哈佛大学和马里兰大学执教，目前是马里兰大学经济系和公共政策学院的杰出大学教授。1991年被选为美国经济学协会会长。谢林在长期从事学术研究与教学之前曾经有过一段在政府部门任职的经历（1945~1953），参与过战后援助欧洲的马歇尔计划，并担任过杜鲁门政府的政策顾问。这些政府经历为他日后思考国际政治和军事策略问题提供了良好的基础和背景。

谢林的获奖可能出乎很多人的意外。按照诺贝尔经济学奖授奖的惯例，诺奖得主通常是那些对经济学基础理论做出重要贡献的人，这种基础性贡献的主要标志就是在标准教科书里能够找到他们的经典模型，在研究生课程的必读著作和论文目录里一定有他们的名字，他们的工作一般是高度技术性的。谢林似乎不属于这样的经济学家。他所有著述较少涉及数学模型，也没有广为引用的经典论文，他的名字只有在人们讨论纳什均衡概念时才会出现，而且是和一本书联系在一起，这本书就是1960年哈佛大学出版社出版的《冲突的战略》。在瑞典皇家科学院颁奖的新闻稿中就特别提到了他的这本书。毫无疑问，《冲突的战略》是谢林获得诺贝尔奖的最重要的原因。

《冲突的战略》有什么特殊之处，使得经济学界在近半个世纪之后赋予它最高的评价？要真正理解这本书在博弈论发展历史上的独创意义和重大影响，我们必须了解当时的一些理论背景。在谢林写作此书的时

候，博弈论尚处于它早期的发展阶段，纳什于 1950 年提出了纳什均衡的概念，并证明了纳什均衡存在的条件。那时博弈论主要作为应用数学的一个分支，对经济学主流的发展几乎没有什么影响。当人们讨论谈判问题时主要从合作性、零和博弈的角度研究什么样的谈判结果是公平和有效率的。谢林没有受当时博弈论发展现状的束缚，而是从非合作、非零和博弈的角度研究冲突和谈判策略的一般理论基础，该理论可以适用于大国间的核战略、国外援助、黑社会帮派冲突甚至大人对小孩的管教策略。更重要的是，谢林不只是满足于一般性的理论探讨，而是直接关注如何从这些理论框架出发深入研究现实生活中的策略互动和选择，尤其是国际事务中的冲突与合作问题。

谢林在《冲突的战略》一书中发展出来的最重要的思想是关于可信承诺（credible commitment）在冲突或谈判过程中的重要作用。这一思想最早源于他的一篇论文“论谈判”（1956 年发表在《美国经济评论》上，后收入本书，成为第 2 章的内容），其基本内涵是，在谈判和冲突的场合，如果博弈一方能够以可信和可观察的方式限制自己的某些选择自由反而会增强其谈判地位，而赋予一方更多的相机决策权则可能伤害该参与人的利益。乍听起来这是不可能的，因为更多的行动自由总是意味着更多的机动性，即便在最差的情况下，你至少可以将机动性弃而不用，保留它怎么会对你不利呢？大家熟知的“置于死地而后生”是这方面最好的一个例子（谢林在书中也提到这个例子）：炸毁唯一逃生的桥梁可以向你的敌人显示誓死一拼的决心，从而达到阻遏敌人进攻的战略目的。在这里，炸毁桥梁就是一种关于绝不退缩的可信承诺，它限制了你的行动空间，但反而收到了吓退敌人的效果。而如果你保有它，给自己留一条退路，看起来似乎万无一失，可是问题在于，这时你再威胁敌人说绝不后退就不那么可信了，敌人可能推断你心虚，进而发动攻击。从这个意义上说，“相机决策”、“因地制宜”这些似乎永远正确的决策原则在一定的博弈环境中并不一定是最优的。

谢林在书中多次提到以合约或立法形式确立可信承诺的意义。最近大陆和台湾两岸政治博弈的进展就提供了这方面的一个绝好的例证。全国人大通过《反国家分裂法》可以看做是大陆方面在两岸谈判僵局中采

取的一种新策略，其意义在于以立法的形式向世界做出一种可信的承诺。这个法案的通过确实限制了大陆在突发事件下的回旋余地，但是另一方面，台湾宣布独立就等同于武力统一的立法承诺消除了人们对台湾宣布独立之后大陆可能采取什么立场的种种猜疑和侥幸心态，使大陆在台湾问题上的谈判地位大大增强了，有利于维护两岸和平。

谢林的思想启发了德国经济学家雷恩哈德·舍尔顿（Reinhard Selten），后者于1965年提出了子博弈精练均衡概念（subgame perfect equilibrium），是对纳什均衡的一种修正和精练，旨在剔除纳什均衡中含有不可信承诺的均衡。舍尔顿正式发展出子博弈精练均衡概念之后，可信承诺这一重要思想才真正在经济学模型中生根发芽，茁壮成长。有趣的是，当我们站在博弈论已经高度成熟和发达的今天，再来重读这本《冲突的战略》，就会发现书中大量关于可信威胁对谈判地位的影响以及如何确立可信的威胁的精辟见解，竟与经济学家二三十年后用严格的博弈论模型所论证的如出一辙，好像谢林当时脑海里已经有了一个精确的数学模型！这种思想的准确性和预见性不得不让人们惊叹。谢林的贡献显然是思想先于模型的典型例证，在这一点上，他和科斯、哈耶克是相似的。舍尔顿的工作也说明了数学模型的力量，思想一旦借助了数学模型就会插上翅膀，展翅高飞。

这里我想举两个例子说明可信承诺在当代经济学中的一些经典运用。诺斯和温加斯特（North & Weingast）的一项著名研究发现，在17世纪英国光荣革命之前，英国王室财政困难，但又无法在民间金融市场上借到钱。富人不敢把钱借给王室，原因很简单，当时的君王拥有无限的权力，不管他如何做出口头承诺，他想赖账谁拿他也没办法。后来新兴的资产阶级和富裕地主通过议会限制王权，尤其是限制随意征税的权力。在君主立宪之后，王室借钱就容易多了。君主立宪限制了君王的权力，是一种“自残”行为，但只有这样，君王借钱还钱的承诺才是可信的。所以说，一个强大的专制政府是无法做出可信的承诺的，专制政府也因其强大和权力无边而受到损害。另外一个例子是科尔奈提出的软预算约束。根据德沃特里彭和马斯金（Dewatripont & Maskin）的新诠释，国有企业的软预算约束问题的根源在于政府无法在企业亏损之后做出不

予救助的可信承诺，因为政府控制了财政资金和银行贷款。只有当政府手中的财力有限，银行的贷款也无法支配或影响的时候，政府对企业不予救助的承诺才可信，企业也就不敢再寄希望于政府了。后来有学者，如安德鲁·沃德（Andrew Walder）、钱颖一和车嘉华教授，在解释中国乡镇企业的崛起时认为，乡镇企业的预算约束比国有企业相对硬一些，因为乡镇政府控制财政资金和银行贷款的能力要相对弱一些。

除了可信承诺之外，《冲突的战略》还包含许多非常精辟和深刻的见解与思想。比如在第 3 章中，他把纳什均衡的结果理解为博弈参与人相互预期对方最可能采取什么策略的“聚焦点”（a focal point），一种双方基于共享经验、文化或认知的默契，为人们理解纳什均衡的选择问题提供了一个有意义的视角。事实上，“聚焦点”理论已成为博弈论讨论参与人如何达成一致或理性预期的一种经典解释。谢林利用“聚焦点”理论解释为什么交战双方有时会自愿达成某种“有限战争”的默契，如不使用毒气和原子弹等。再比如谢林提出，在动态博弈中，拥有更多信息也许是一件坏事，不知道反而更好。谢林提到一个有趣的例子，一对夫妇在电话里为碰面的时间与地点争执不下，如果此时妻子说完何时何地碰面之后立即把电话挂了，那丈夫也只能听从妻子的意见。小孩有时会故意大声哭叫以装着听不见大人的呵斥和警告。在很多场合，策略地屏蔽信息、变得“无知”是增强自己谈判地位的一条捷径。我们经常说的“眼不见心不烦”“无知者无畏”与谢林的这一观点有相通之处。书中诸如此类的真知灼见不胜枚举，读者可细心体会。

谢林的研究领域极为广泛，从核扩散、军控、恐怖主义、有组织犯罪到能源和环境政策及种族隔离问题，著述甚广。他另外一本影响较大的书是《微观动机与宏观行为》（Mircomotives and Macrobehavior），1978年由诺顿公司出版，研究在社会领域个体的选择如何导致了社会层面的系统现象，如种族隔离和性别歧视。谢林自己也承认，对于长达 50 余年的研究生涯，很难用一种简洁的语言概括自己的研究兴趣究竟是什么，但在他自已看来，这一切研究都相互关联，而其中最重要的纽带是痴迷于理解人们（包括国家）如何相互应对和影响对方。确实，从博弈论的角度研究人们之间的互动及其影响，构成了谢林所有学术贡献

的一条主线。

博弈论对于经济学的重要贡献使得三位经济学家——纳什、舍尔顿和哈桑尼——分享了1994年诺贝尔经济学奖的殊荣。时隔11年之后谢林和奥曼又因对博弈论的巨大贡献而获奖，这凸显了博弈论在当代经济学中的核心地位。然而，这些经济学大师的原创贡献大都体现为高度技术性的论文，让一般的读者望而却步。所幸谢林的原创贡献是一个例外，他的《冲突的战略》没有涉及多少专业术语和数学模型，完全依靠逻辑的力量展现思想的深度和广度，无疑是体现博弈论如何帮助人们理解现实世界的最好范例。我相信，任何对博弈和策略感兴趣的读者都将从本书的阅读中直接获益，因为智慧在深层次上是相通的。

华夏出版社的编辑陈小兰博士从去年上半年开始与我商议出版一套当代经济学经典译丛，在我开出的书单里第一本就是谢林的《冲突的战略》。经过译者的精心努力，译稿于今年7月份完成。后又承蒙王勇博士和郑志刚博士的热情协助，他们对译稿进行了全面仔细的校对。在此书的编辑出版之际得知谢林因为此书获得了诺贝尔奖。古人云：“谋事在人，成事在天”，我想，天底下没有任何促销广告能比这个突如其来的喜讯更具广告效应了。是为序。

周黎安
北京大学光华管理学院
2005年11月12日

1980年版序

闻知哈佛大学出版社将要出版《冲突的战略》平装本，我顿感自己有必要删除或重写书中一些陈旧的部分。或者至少，我应该在该书新序中向他们表示歉意。虽然《冲突的战略》一书已经出版了 20 多年，但是我平时很少完整地读过此书。书中有些部分，我甚至 10 多年都没有看过了。我必须指出，书中有些部分已经陈腐，或者不切主题，更有甚至是完全错误的。

当然，这只是就部分内容而言。总体来说，我认为，尽管书中有些案例看似古怪离奇，但是该书大部分内容还是正确的。现在看来，我在该书第 1 章中有关军事战略不被大学学者和军事部门重视的论述存在明显错误。这些内容除了其历史价值外，可以完全丢弃。更严重的问题在于学生，特别是那些第一次阅读《冲突的战略》的学生，是否熟悉书中出现的赫鲁晓夫和穆沙德（Mossadeq）等历史人物以及瑞恩高德（Rheingold）小姐的当选历程。

尽管如此，令我们感到欣慰的是该书附录 A 的内容尚未过时，其存在前提是原子武器自美国在日本长崎首次使用以来从未被再次用于战争。希望《冲突的战略》再版之时，该前提依然成立。

最初我在该书第 10 章中提到的一些观点，今天已经广为人们所认可。不可否认，其中有些观点已经过时。现在存在很多有关军控的文章，其中许多观点我都曾经论述过。而且，第 10 章中涉及很多限制战略武器条约（*the Strategic Arms Limitation Treaty*）的部分与我现在找到的其他 25 页相关文章论述得一样准确。

希望进一步了解本人有关战略和军控思想的读者，可以查阅我和莫顿·H.哈尔伯林（Morton H.Halperlin）合著的《战略与军控》（*Strategy*

and Arms Control）一书，该书由 21 世纪基金会于 1961 年出版；或者参阅我编著的《军备与影响》（*Arms and Influence*），该书由耶鲁大学出版社于 1966 年出版。

现在，大多数人阅读《冲突的战略》一书是为了满足理论需要，而非实际外交决策。我之所以收集整理众多资料编写此书，主要目的是为了建立一个跨学科领域——该领域曾被人们称之为“讨价还价理论”、“冲突理论”或“策略理论”。我想向人们证明，一些横跨经济学、社会学和政治学，甚至是法学和哲学或者是人类学的基础理论不仅对从事理论研究的学者大有裨益，而且也有助于人们解决实际问题。与此同时，我最初还认为博弈论（the theory of games）在某些领域应该具有更多实用性。但是现在看来，这也许是一种曲解，因为除了大家比较熟悉的霍华德·雷夫（Howard Raiffa）、马丁·舒里克（Martin Shurik）和纳杰尔·霍华德（Nigal Howard）等学者外，博弈理论家更多倾向于将博弈论限制于数学范畴。尽管我致力研究的领域可能不会突飞猛进，甚至尚未有自己的正式名称，但是我相信将会有更多的有志之士投入到这一领域的完善和发展之中。

现在，一些期刊杂志也为这一领域的发展做出了重要贡献，其中尤其值得一提的是《冲突解决杂志》（*Journal of Conflict Resolution*），这些期刊为了满足实际决策者以及实际工作者等非专业人士的需要，除了一些如没有过多涉及博弈的基本概念，等博弈术语外，甚至是某些基本理论。[仅仅在几年前，为了更好地帮助读者理解文章内容，我在一篇讨论美苏两国对军控武器态度的文章中引入了一个 2×2 的矩阵图例。但是，杂志社的主编（在此我不便透露其姓名）坚持让我将其中的矩阵图例删掉以免吓着读者，并表示文字说明已经足以满足文章的需要，尽管他本人对我的意图也没有完全理解。]

自出版以来，《冲突的战略》受到了读者的广泛好评。许多读者向我反映，他们非常喜欢这本书并从中受益颇深，这令我感到由衷高兴。但是这 20 年来，最令我感到欣慰的是后来约翰·斯特雷奇（John Strachey）对该书的评价。约翰·斯特雷奇是 20 世纪 30 年代杰出的马克思主义经

济学家，我在大学时曾拜读过他的著作。第二次世界大战后，他担任英国工党执政政府的国防大臣。因为他写过一本关于裁军和军控的书，哈佛大学国际关系研究中心许多学者曾邀请他来哈佛进行学术访问。当约翰·斯特雷奇来我家做客时，他高呼《冲突的战略》对他的研究帮助非常大，并兴高采烈地滔滔不绝讲个没完，而此时我却在思考该书到底是哪个章节的哪些观点对他产生了如此大的影响。事实表明，并不是因为某个章节的某个观点与他的观点有分歧，而是直到读了《冲突的战略》一书，约翰·斯特雷奇似乎才领悟到非零和冲突的存在。此前，他只知道冲突能够与共同利益并存，并想当然地认为二者存在本质上的区别，而不是单一整个结构中的两个部分。这位前半生致力于垄断资本主义与阶级冲突、核战略与联盟政治研究，而在其学术生涯后期却研究军控和维和的学者读了《冲突的战略》后，被书中的精辟分析——甚至连我本人都未意识到——所震惊，并谦虚、郑重地把此事告诉我。由此可见，真的很难说一本书会产生什么影响。

托马斯·谢林

于马萨诸塞州剑桥

第一部分

战略理论构成要素

第1章　缓慢发展的国际战略学
第2章　论谈判
第3章　谈判，沟通与有限战争

第1章

缓慢发展的国际战略学

在研究“冲突”的众多理论中，对于“冲突”一词大体存在两种不同的解释：一种认为冲突是一种不正常（pathological）状态，并寻找产生冲突的根源和解决冲突的方法；另一种认为冲突的产生具有合理性，并研究分析与冲突相关的各种行为。后者可进一步分为两派：一派主张对冲突主体进行综合全面的分析，包括冲突主体的“理性”与“非理性”、有意识与无意识的行为及其动机（motivations）和对利弊的权衡（calculations）；另一派则更关注冲突主体充满理性、意识和智谋的行为。可以说，后一派把冲突看做是一场冲突双方都“志在必得”的竞赛。研究这些有意识的、理智的、复杂的冲突行为，特别是冲突中获胜一方的行为，如同寻找赢得竞赛的“正确”行为法则。

我们可以把这一研究领域称为“冲突的战略”[1]。至少有三个理由足以激发我们对这一领域的兴趣：我们自己可能身处冲突之中；事实上，我们每个人都不同程度地卷入国际冲突之中，并“志在必胜”。我

[1] 这里的“战略”一词取自《博弈论》一书。该书严格区分了技巧性博弈（games of skill）、机会博弈（games of chance）和战略博弈（games of strategy）。在战略博弈中，博弈一方的最佳对策取决于另一方的行为。“战略”一词强调博弈双方决策的相互依赖性以及双方对彼此行为的预期判断。这里的“战略”不同于军事意义上的“战略”。

们也希望知道冲突主体在现实冲突中的真实反应；对“正确”行为法则的理解为我们提供一个研究现实冲突行为的基点。如果我们希望控制或影响冲突主体的行为，就必须首先弄清楚控制过程中的各种变量如何影响冲突主体的行为。

如果把研究仅仅局限于冲突论，我们将受到理性行为假设的严格限制，它不仅指明智行为，还包括权衡利弊后所做出的行为反应，而这种权衡则建立在冲突双方各自明晰的、内部协调的价值体系基础之上。这样的话，我们就人为地缩小了研究结论的适应范围。如果仅仅为了研究现实中的冲突行为，我们在这种局限的条件下得到的分析结论要么是对现实完美的反映，要么则是对现实的歪曲。由于任何抽象活动都存在类似的风险，因此我们必须做好对自己的研究结论进行评判的准备。

有关“战略”的研究之所以能够推动理论的发展，并非这一研究方法在目前所有的研究方法中最能反映事实，而在于其对理性行为的假设卓有成效。这种研究方法能够一针见血地揭露研究对象的关键所在，从而有助于理论发展。此外，这种方法能够使我们严格区分自己分析的冲突过程和假设中冲突主体参与的冲突过程。我们要求假设中的冲突主体行为具有某种连贯性，因为只有这样我们才能根据他们是否符合连贯性这一标准来分析是否存在其他可供选择的行为模式。虽然对“理性行为”的假设极大地推动了理论进步，但是我必须重申，检验研究得出的理论是否真实反映了现实则是后续之事。

虽然战略论承认冲突产生的合理性，并以假设的“志在必胜”的冲突主体作为研究对象，但是并不否认冲突双方之间除了冲突利益往往还存在某种共同利益。事实上，研究对象的多样性本身反映了国际关系中对立冲突与合作依赖并存这一现实。双方利益完全对立的完全（pure）冲突状态是非常罕见的。完全冲突通常只会在大规模毁灭性战争中出现，否则在一般战争中也很难发生。鉴于此，冲突中的“胜利”一词含义并没有严格的界定。冲突中的“胜利”只是相对于冲突一方的价值观而言，而非冲突另一方。而且这种胜利只能通过冲突双方的讨价还价（bargaining）、互谅以及避免采取互损行为得以实现。如果战争成为解

决问题的惟一方式，那么就会出现所谓的“完全冲突”。但是，倘若存在任何避免发生大规模毁灭性战争；或是仅打一场破坏性最小的有限战争；或是武力威慑就能迫使对方退缩的可能性，那么互谅的可能性就同冲突要素一样重要。诸如在威慑、有限战争、裁军或谈判中，冲突双方都隐含着某种共同利益和相互依赖性。

因此，我这里使用的“战略”一词仅涉及武力威慑（potential force）而非武力的有效使用。另外，此处所指的冲突双方也不是完全誓不两立的敌人，而是互存疑虑与分歧的伙伴。这里关注的不是冲突双方的利弊得失，而是如何达到一个实现冲突双方最大利益的结果。在博弈论的术语中，最有趣的国际冲突不是“零和博弈”而是“非零和博弈”。在后者，博弈一方的所失并不意味着另一方的所得；对博弈双方而言，寻求一个双赢结果就是他们的共同利益。

我们研究分析冲突的战略时，必须看到大多数的冲突都存在讨价还价的可能性。在这种情况下，冲突一方能否达到自己的目的取决于另一方选择或决策的最佳平衡点。如果冲突一方率先妥协，双方讨价还价的可能性就变得十分明显；或者当冲突一方完全占领或退出具有战略意义的领土后，讨价还价就成为一种默认行为。这如同市场中常见的讨价还价行为，找到双方利益的最佳结合点，做出对双方互利的安排。但是，其中也可能存在如罢工、联合抵制、价格战或强取豪夺等破坏性行为，甚至造成两败俱伤。

将冲突视为一个谈判过程，有助于我们摆脱非敌即友的传统思维模式。我们把有限战争看做是一个讨价还价的过程，意在强调冲突双方之间除了存在利益分歧外，寻求双赢结果也是共同利益之所在。成功的雇员大罢工绝不是搞得雇主饱受经济损失、资不抵债，否则根本就不存在所谓的成功罢工。战争也是同样的道理。

“威慑”思想经历了一个漫长的发展历程，这无疑对我们的研究工作十分有利。在“威慑”成为我们国家核心战略的十多年里，“威慑”思想不断被完善和发展。我们知道，威慑必须可信、有效，威慑的可信度在一定程度上取决于执政党实施威慑的成本和风险。长期以来，我们主张实施威慑，建立遏制敌人进攻的包围圈，并使其发展成为一种民族

荣耀和尊严。这一切都增强了威慑的可信度，从而推动了“威慑”思想不断发展。我们认识到，一旦发生突发事件，我们可以避重就轻，通过在局部地区发动有限战争，避免大规模的报复行动。与此同时，研究发现，如果诸如不久前提出“核共享”（nuclear sharing）计划的那些顽固之徒拥有足够的资源发动大规模报复行动，并有能力为此承担一切后果，那么他们就可能发动大规模报复行动。此外，我们还注意到，对手的自我控制力与其实施威慑的有效性存在密切联系。威慑对那些丧失理智、行若稚童的人不起任何作用。我们的研究还发现，威慑的效果与对手掌握的可供选择的资源有很大联系。如果不是一头困狮的话，对方的行为多少令我们能够容忍。但是我们必须承认，孤注一掷的威慑则令对手更加危险，因为在这种情况下，他别无选择，只能与我们殊死一搏。这样也就排除了任何回旋的余地，迫使对方必须在两个极端之间做出抉择。研究表明，大规模杀伤性威慑只有当对方完全不具备杀伤性报复能力时才能发挥作用。因此，我们必须考虑对方的突袭能力是否达到足以导致对方先发制人，以避免首先遭受来自我们的突袭。最近，我们开始研究所谓的“警惕突袭措施”问题，考虑是否能够通过军备控制缓和彼此间的威慑。

令人感到震惊的不是演变中的威慑思想变得多么复杂以及人们如何小心翼翼地发展和完善威慑思想，而是威慑思想的发展步伐如此缓慢，威慑概念如此含糊不清以及现存的威慑理论体系如此松散。当然，我这样说并不是为了打击那些十多年来一直致力于发展威慑概念的学者专家。关于威慑的相关战略问题就是一个范例。那些制定具体政策、解决紧急问题的决策人员并没有或者很少从现存的理论体系中得到某些启示和帮助，从而不得不在决策过程中创造他们自己的理论依据。至今，甚至尚无相关的威慑科学读物能与诸如通货膨胀、亚洲流感、基础教育和烟雾等方面的读物相媲美。

此外，大多数人研究战略思想是为了解决实际问题，而非真正关心这一理论体系的完善和发展。不仅从事实际决策的政府人员和媒体记者们抱着这种想法，一些学者亦如此。无论是否反映了一些学者或编者的利益需求，有关威慑及其概念的著作文章主要是为了解决实际问题，而

非发展解决问题的方法论。[1]正因如此，我们现在甚至没有自己的专业术语，偶尔使用的“积极”（active）威慑和“消极”（passive）威慑等词汇并不能满足理论发展的需要。

那么，我们其中谁应该为威慑理论的缓慢发展承担责任呢？一个严峻的现实引起了我的思考。与其他几乎所有庞大的、受人尊敬的行业相比，军事部门似乎还没有相对应的学术研究机构。在经济、医学、公共卫生、国土资源保护、教育以及刑法等行业部门，它们都有大量相对应的学术研究机构。（以经济部门为例，从事经济研究和著作研究的人员数量几乎等同于从事实际经济决策或行政管理人员的数量。）但是我们军事部门的情况又是如何呢？

我们没有相当规模的军事研究机构，现在的大学也大多只是从事教学工作，而非理论研究，更不用说相当规模的军校或者军事部门中建立的其他非技术性教研机构了。这些院校或者研究机构缺乏长期稳定的研究人员、明确的研究方向以及持续性系统理论研究所必备的价值观。

在美国大学里，仅有一些为数不多的历史学家和政治学家从事国家战略研究。他们的存在价值仅仅在于向相关政府部门提供建议方案，遏制苏联占领欧洲；其重要性最多停留在实施“反垄断法”的水平上。我之所以这么说，没有任何贬低现有研究成果的意图，而是为了强调美国大学通常还没有设立相关的咨询机构或热线，与军方或者从事外交事务的权力人物存在任何密切联系。［最近制定实施的“后备军官训练队”(ROTC) 计划也仅仅是一个小小的例外。但是，至少该计划涉及了能够提供历史学和政治学等相关课程的机构。］现在，美国许多大学都设立了国防研究项目或相关机构；许多基金会也开始关注国家安全问题，这

[1] 当然现实中也存在很多这方面的反面例子，如 C.W.Sherwin,Securing Peace Through Military Technology,Bulletin of the Atomic Scientists,12:159-164 (May,1956)。该文引用的沃伦·阿姆斯特 (Warren Amster) 的一篇论文提醒我们，当理论发展仅仅为了满足军事需要时，如同现在的情况，那么，理论研究结论就不可能被公开发表。毫无疑问，其中也有来自媒体编导方面的阻碍。国际关系刊物大多面向非理论研究的普通读者，高度理论性的文章当然会被拒之门外，而且这些刊物关注的大多是热点问题。参见 Anatol Rapoport,Lewis F.Richardson’s Mathematical Theory of War,vol. Ⅰ,No.3 (September 1957)，该文是对冲突解决方法研究的重大贡献。

无疑是一个全新的重大进步。新型的准政府研究机构，如美国兰德公司（RAND）和国防分析研究所（the Institute for Defense Analysis）为弥补相关空缺发挥了举足轻重的作用。这显然符合我们的初衷。

也许有人会问，难道军事部门没有能力提出它们自己的相关理论，以丰富威慑和有限战争的思想吗？毕竟，理论的发展并不仅仅是大学里那些与世隔绝的专家学者们的事情。如果军方具有足够的智慧能有效地运用（*use*）军事力量，可以说他们已经具备将其理论化的能力。但是武力的使用（*Application*）和武力的威胁（*threat*）之间存在很大的差异，这对我们来说是非常有利的。威慑涉及武力的潜在使用，而非实际使用；威慑是为了警告并说服潜在敌人，只有避免某些举动才能维护其自身利益。因此，虽然都是为了追求国家目标，但是执行军事任务与运用潜在（*potential*）的军事能力二者所要求的智慧存在极大的差异。本质上讲，威慑理论是一门关于如何巧妙地运用智慧，避免使用现实武力的理论。鉴于此，威慑需要的不仅仅是军事策略，其含义非常广泛。军人具备的能力可能更广、更强，但是他们不会自觉地表现出来，因为他们有自己的天职，而履行天职是他们的第一要务。[1]

15年前，博弈论的一个新发现促使了战略论的诞生。博弈论研究的核心问题——“战略”博弈、“技巧性”博弈和“机会”博弈——是博弈一方的最佳策略取决于其对另一方行为的先期预测。威慑恰好满足这一定义的条件。威慑思想之所以能够发挥作用，不仅仅在于对手期望我们对其举动做出相应的行为反应；同时，我们之所以能够制造威慑，因为我们期望借此对对方的选择施加影响。但是迄今为止，现实的国际战略环境还不能完全满足博弈论的前提假设条件。事实证明，博弈论有

[1] 伯纳德·布罗迪（Bernard Brodie）在其著作 Strategy in the Missile Age (Princeton, 1959) 中深入探讨了为什么在军事战略领域缺乏充满活力的学术传统。相关的文章有 Colonel Joseph I. Greene’s 为 Clausewitz 的著作现代版 On War (New York,1943) 写的序中写道：“在大国争战的年代，为了加快学员参战的步伐，我们的两所一流军校不得不把课程缩短到 10个月……因而，无论是从时间上还是从学术环境上，都不允许他们进行长期的军事思想研究……如果能给我们的学员留下更长的训练时间，哪怕是 2~3 年的时间，我相信军事思想家们一定会有属于他们自己的研究成果。”

助于人们将国际政治问题公式化以及对相关概念的诠释；但是博弈论最大的成功之处在于其他领域，而非国际战略理论领域。总体而言，博弈论倾向于抽象的理论层面，而且即使在这一层面，博弈论也很少涉及诸如威慑这样的现实问题。[1]

尽管不是整个国际关系领域，威慑思想在冲突理论领域某些方面的作用也十分显著，以致有人认为在国际关系领域存在一个有待于人们发现，并被用来指导国际关系实际的完整理论体系。很长一段时期以来，威慑都是作为刑法中一个非常重要的概念而存在。数十年来，人们一直期望立法者、执法者、律师和法学专家们能够对威慑概念进行严谨、系统的研究论证。毫无疑问，威慑概念并不是刑法中惟一需要考虑的事项，甚至并非不可或缺的最重要因素。尽管如此，在一些人看来，威慑思想还是为研究有关问题提供了一个可供参考的理论依据，如法官对既定罪犯量刑的尺度、嫌疑犯的价值观念、犯罪行为的危害性、现存法律体系对罪犯行为的解释力、罪犯的法律意识、故意犯罪行为的多样性、社会对犯罪行为惩罚成本的承受程度以及罪犯对此了解的程度、体制弊端的可能性、第三者假公济私的可能性、沟通在组织社会与罪犯之间的作用、犯罪组织对社会体系的威胁等等。

威慑不仅适用于犯罪分子，也适用于我们自己的孩子。威慑思想甚至在某些方面借用了家长教育孩子的规则，例如教育威慑客体保持自律和理智的重要性；威慑客体如在黑暗和噪音状态中得知危险，对危险释义的能力的重要性；如果需要，威慑主体实施威慑的意志，特别是威慑客体确认威慑存在的能力的重要性。但是，对犯罪分子的威慑与对孩子的威慑最大的不同点在于，家长在惩罚孩子的同时，也伤害了自己。此外，父母对孩子的威慑更类似于一个富有的老子国（paternalistic nation）威慑一个贫弱的、政府组织混乱的国家，即：前者强迫要求以提

[1] 杰西·伯纳德（Jessie Bernard）在其文章 The Theory of Games as a Modern Sociology of Conflict 中对此给予了充分肯定。他强调：“我们可以预见，数学家们在不久的将来会把博弈论用于社会现象的分析。”（The American Journal of Sociology,59:418,March 1954）我个人认为，现存的缺陷不在数学问题，而在于战略理论已经深受社会学家的影响。他们很大程度上，把战略理论看做或认为其应该是数学的一个分支。

供外债和援助换取后者“优惠”的经济政策和军事合作。

这一比喻事例告诉我们，即使在国际关系中，友好国家之间和敌对国家之间都存在威慑的可能性。［例如美国威胁法国，如果法国拒绝批准《欧洲集体安全条约》（European Defense Community Treaty），美国将实施“地缘战略”（peripheral strategy）。实际上，这样的威胁如同报复威胁一样毫无作用。］威慑存在的前提是双方存在冲突的同时，也存在共同利益。如果双方之间的关系是完全共同利益关系或者完全冲突关系，那么，威慑无法发挥其应有的作用。在上述两个极端之间，对盟友的威慑和对敌国的威慑之间的区别也只是程度不同而已。如果我们希望负责任地指出，美国与苏俄或希腊之间是否存在相对于分歧而言的共同利益，[1] 我们应该首先建构一个富有连贯性的理论体系。

在现实生活中，我们也可以随处看到威慑思想的身影。汽车司机之间也存在避免发生撞车的共同利益和谁首先停车让路的冲突。对双方司机而言，撞车是双向的、也是惟一可以使用的威慑手段，任何一方都有权以撞车报复另一方侵犯自己的通行权。这个案例启发我们：实际行动可以和语言一样产生威慑作用；威慑不仅可以表现为对外声明，也可以是威慑能力的丧失。

最后，我们必须提到黑社会中一个非常重要的层面。黑社会内的帮派斗争与国家间战争有很多相似之处。国家之间与黑社会帮派之间都没有强有力的法律体系调整他们之间的关系，二者最终都将付诸武力。虽然他们之间都存在避免冲突的共同利益，但是又都喜欢采取暴力行动威胁对方。一个非常有趣的现象是，诈骗犯和犯罪团伙之间也存在所谓的有限战争、裁军、脱离军事接触、突袭、报复和报复威胁。他们也担心被招安（appeasement）或失去面子。与国家一样，他们也喜欢彼此结盟，缔结一些条约。但是，这些条约毫无法律保证，因为一旦一方违

[1] 在此有必要强调，我这里所指的“共同利益”并不是通常所说的价值观念上存在共同性，而是他们之间是同乘一条船的关系。他们或许仅是为了占据战略优势而身在该处——防止翻船是他们的共同利益。假设二人同乘的船沉了，给他们一组选择，他们之间就存在我这里所指的“共同利益”。“潜在的共同利益”似乎更具有描述性。例如，威慑就是将一方的行为过程与另一方的行为过程结合起来，寻求共同利益。

约，另一方根本没有倾诉之处。

事实上，现实中存在的众多研究领域也都将为我们关注的国际关系研究提供某些帮助和启示。但是很多时候，在我们关注的领域中，一个原则或者隐藏在众多细节背后，或者因其本身可能是一个复杂、庞大的结构体系，或者因为我们假设的前提错误而没有被发现。然而，在一个结构简单易懂的研究领域或我们前提假设正确的领域内，这个原则更容易被发现。当某个人徒劳无功地想通过威慑阻止一个孩子伤害一条狗或阻止一条小狗伤害孩子时，我们也许能够更容易准确地描述使用威慑遏制一个穆沙德的异常困难性。

似乎尚未有任何冲突研究领域建构起了一个略加修饰即可适用于国际关系分析的完整理论体系。包括研究黑社会冲突中犯罪行为的社会学家，习惯上也从未认真关注过我们这里所说的冲突的战略（strategy）。即使那些研究法学和犯罪学的论著也没有建立一个完整明确的相关理论体系。我不敢说黑社会中是否存在关于勒索理论的单行本、书册或原始手稿。但是有一点可以肯定，除了《少儿教育新编》这样涉及“威慑”的书籍外——尽管呼声很高——尚不存在有关如何使用或对抗讹诈的类似修订版论著[1]。

战略学理论应该包括哪些要素？这一理论应该回答什么问题？该理论体系应该统一、明确、有效地向人们传达哪些信息？首先，我们必须为问题中所涉及的有关场景和行为要素进行定性。威慑——这里指典型的战略概念——是指一方采取有效方式影响对方决策，并期望借此影响对方对自身行为预期判断的行为模式。威慑涉及与对方冲突的同时，以足够的证据使对方相信我们的反应取决于对方的行为。

但是，威慑双方间什么样的价值观念结构——即博弈论中的“决定因素”——才能让威慑更好地发挥作用呢？我们怎样计算构成“威慑”条件的冲突和共同利益的分配比例呢？需要向人们传达什么信息，以及以何种方式传达？威慑客体应该具备哪些“理性”（rationality）条

[1] 尽管如此，相关理论研究还是取得了一定的成果。其中包括 Daniel Ellsberg inchuded a lecture on “The Theoy and Practice of Blackmail,” and one on “The Political Uses of Madness,” in his series on “The Art of Coercion,” sponsored by the Lowell lnstitute,Boston,March 1959。

件——是对自我价值观的释义、辨析有利条件和权衡各种可能性的能力，还是展现（或无法隐藏）自我理性的能力？

实现互信或履行承诺的条件是什么？具体地说，就是除了威慑已经造成的损失之外，如果一方开始履行承诺，是否需要另一方保证停止进行中的破坏行为？或者说，这是否取决于双方间决定因素的比例结构？什么样的“司法体系”、沟通体系或信息结构是承诺成立的必要条件？

一方能否宣称自己“可能”采取威慑行动；或者一方宣称自己将一定采取威慑行动？如果一方在宣称自己“可能”采取威慑行动后，又持保留态度，无意采取相关行动，那么他此前宣称的威慑又有什么意义呢？一般而言，如果一方考虑到已经做出的威慑承诺没有实施即已产生了预期效果，那么，他会采取什么策略来处理自己已经做出的却又不愿意履行的威慑承诺？如果存在的话，一个能够阻止对方行为的威慑与一个促使对方行动的威慑之间，或者与旨在通过自我的人为失误来保护另一方的威慑之间存在什么区别呢？威慑、惩戒以及敲竹杠式的威胁之间是否存在逻辑上的不同？

如果第三方与冲突双方存在冲突利益和共同利益，能够接触或控制沟通系统，并且采取一定意义上的理性行为（或非理性行为），而且与二者之一存在互信或有效的协议，那么，第三方如何影响威慑局面呢？如果存在一个允许或禁止某些行为的司法体系，不仅有权惩罚不履行承诺的一方，并且有权要求双方提供可靠的信息，那么，这个司法体系如何对这些问题产生影响呢？考虑到现实中的或假设的司法体系，或者选手的价值观念的变化，或选手与现实中（或假设）的第三方关系，我们需要将“荣誉”、“脸面”或“信任”等概念理性化到何种程度？

上面这些简短的问题表明，建构战略“理论”的空间还很大。看起来，这一理论似乎融合了博弈论、组织论、信息论、实证主义以及机会主义和集体决策论。这也符合我们对“战略”一词的定义条件：既重视冲突的存在，又关注冲突双方之间的共同利益；既重视“合理”追求价值最大化的行为模式，又关注一方的“最佳”选择取决于另一方的行为，以及“战略行为”涉及的一方如何通过研究自我对对方行为的预期判断来影响对方决策这一事实。

这里需要强调两点：一方面，尽管“冲突的战略”这个名字听起来让人毛骨悚然，但是这一理论既不涉及武力的有效“使用”，也不主张先发制人或以暴制暴。相反，“冲突的战略”理论主张运用威慑，或威慑与承诺并存，或者更通俗的说法就是一方行为对另一方行为的影响。这也是该理论的核心所在。

另一方面，“冲突的战略”理论对冲突与共同利益一视同仁，不仅适用于潜在敌人，也适用于潜在朋友。这一理论只有在下面两个极端条件下才失去其自身作用，即：在双方之间存在完全的共同利益，即使在寻求和达到共同利益的过程中也没有出现任何分歧；或者二者之间存在完全的冲突，即使发动两败俱伤式的战争也没有丝毫共同利益可言。即便如此，在上述两个极端之间，“冲突的战略”理论也没有严格划分冲突与共同利益的界限。我们既可以称其为动态（precarious）理论，也可以称其为不完全对抗（incomplete antagonism）理论。[1]（我在第9章中指出，国际关系中有关突袭的一些重要问题都与双方之间的互不信任问题存在结构上的相似性。）

《冲突的战略》理论对于双方冲突程度的中立态度，以及“战略”一词涉及通过影响对方对自我行为后果的预期判断达到遏制对方的目的表明，我们完全有理由把我们所研究的领域称为相互依存决策理论（*the theory of interdependent decision*）。

威慑与反威慑、报复与反报复、有限战争、军备竞赛、边缘政策、突袭、攻击与欺诈等一系列行为既可以看做是理性行为，也可看做是一时的非理性行为。尽管为了理论发展的需要，这些行为都被看做是理性行为，但是实际并非如此，有些仅是一时的头脑发热。当然，假设上述行为是理性行为，并不意味着这些假设一定有助于系统理论的发展。如果实际中的行为都是理性行为，建构相关的有效理论就会比实际更容易。假设我们把自己的研究结论看做只是评判现实的基点，而非有效的

[1] 我这里使用“威胁”一词，没有任何主张进攻或敌对的意思。在朋友之间开诚布公的谈判或默契合作中，任何分歧和影响合作的消极因素，无论是公开的还是暗示的，都是双方支持各自要求的手段。这类似于商务谈判中，供不应求通常会导致价格攀升。

完整理论的话，我们应该努力保护自己避免遭受一个存在缺陷的理论导致的最坏结果所造成的不利影响。

此外，鉴于这一理论成立的前提是冲突双方能够保持冷静，理性地根据各自一贯的价值观念权衡利弊，我们不得不对“非理性”(irrationality）的含义进行更加全面、深入的分析，因为现实中的决策者们不是简单地分布在连接理性与非理性的一维层面上。理性是一个人个性特征的总体反映，除了完全理性外，还存在其他众多构成要素。非理性意味着无序和非连贯性的价值体系、错误的分析判断以及无法有效地获取信息与沟通；非理性还意味着在接受或传达指令、获取或传递信息过程中可能存在危险的随机因素。此外，非理性有时仅仅能够反映众多个体决策的集体特征。但是事实上，由于每个个体具有不同的价值观、组织形式和沟通体系，从而根本无法像单一实体那样产生统一的行为模式。

实际上，理性行为模式通常由众多关键要素构成，而且这些共同要素又与某个具体的理性或非理性行为体的行为存在极大相似性。价值观念、沟通体系、信息系统、集体决策过程，甚至表示错误或失控存在的参数都可以被看做是对非理性研究模式化的努力。希特勒、法国议员、轰炸机飞行员、珍珠港的雷达操作员、赫鲁晓夫以及美国选民可能都有某种非理性特征，但毫无疑问，他们又不可能是同一个人。其中有些人的行为可以被看做符合理性行为标准。（甚至不具有稳定价值观和无法令自己安定的精神病人，都乐于选择压抑而非鼓励自相冲突的目标。因而从某种意义上讲，精神病人也被认为具有某种理性人的特征，即：具有特征鲜明的价值观念，通过具有随意性、非对称性沟通、危险性等特征的选择过程实现集体决策。）

另外两个研究结论削弱了理性行为——如利弊权衡、收益最大化战略——本身所具有的局限性。第一个研究结论是，甚至在那些经济条件不平衡、被确认为非理性的群体中也存在对战略与生俱来的倾向性，至少存在对战略原则的运用。当然，在此我必须声明，这并非我本人的研究结论。我被告知，即使那些精神病院里的病人也似乎懂得如何有意识、如何本能地培养自己的价值观，从而能够向他人实施威慑或保护自我免遭他人的威胁。对待伤害，一个草率的甚至具有自我摧残性的态度

就是“如果你不让我……，我就割断自己的静脉”。这也是人类本身所具备的一个先天性战略优势。诸如此类还有听觉和理解力的丧失，或者能够破坏外界惩罚性威胁或威慑的经常性自我失控等（这令我再次想起了自己的孩子）。事实上，在冲突与共同利益并存的条件下，理性战略决策理论的优势之一在于，这一理论能够通过揭示看似自相矛盾的谈判技巧的本质，向大家展示弱势人群同样能够演绎某些理性的谈判技巧。如果说我们后天形成的圆滑世故某种程度上摧残了自我天性，这也许有些言过其实。但是，一个简单的理论结论告诉我们应该反璞归真，重新找回我们失去的那些看似“非理性”的天性。

第二个研究结论与前一个密切联系。有关“理性”决策及其产生战略性影响的理论研究恰如其分地表明，在冲突不可避免、决策和动机看似理性的情况下，“理性”行为不是一个放之四海而皆准的有利条件。正如前面所提到的，构成理性的众多要素中的很大部分是冲突条件下的战略不作为（strategic disabilities)。希望自我非理性或者——如果这种语言在哲学意义上存在争议——期望得到在特定条件下避免某种理性行为的权力都可以被看做是一种理性行为。个体可以放弃和破坏自我理性能力，至少可以限制到某种有限程度。他之所以能做到这些，因为构成理性的各个要素并不仅仅是人类精神不可分割、深度个性化以及完整性的属性，而且包括人类听觉辅助功能、邮件的可靠性、司法体系及个体代理人或伙伴的理性等因素。原则上，逃避威胁的方式多种多样，如麻醉大脑、隐退、被没收财产、打断自己签支票的手臂等。在战略理论中，如果非要给这些行为定性的话，我们可以把消极抵抗行为看做是对理性的削弱。对于一个把“理性”简单看做基本原理的理论，我们不仅无法完善和检验该原理，更无法从中找到任何隐藏的规律。事实上，在冲突条件下，“理性”看似矛盾的作用本身就证明了系统理论能够提供有力的支持和帮助。

尽管如此，对战略行为理论分析得到的结果看起来仍似乎前后矛盾。它们往往与常识或已经被普遍接受的规则相矛盾。正如前面的威慑案例提到的，面对威胁时，特别是理性与非理性同时存在时，保持理性是绝对正确的论断似乎存在很大漏洞。当面对威胁时，具有良好的沟通

体系、完备的信息系统，以及完全掌控自己的行动和财产也许并不是什么好事，对待穆沙德和教育孩子就是典型例子。同样的情况还有截断后路、背水一战，告诉敌人自己已经无路可逃。根据战略论，一条规定供奉海盗将受到严厉惩罚的英国老式法律条文并不是冷酷无情或一反常态的。有趣的是，政治民主本身也必须依靠某种沟通体系，这种体系能够消除确凿证据的传播渠道：强制使用匿名投票箱是为了杜绝投票人透露其投票方式；被剥夺表达自己选举方式的权利也就意味着失去了被威胁的权利。既然投票人无法对外证明自己是否受到威胁，投票人和实施威胁的人都知道任何惩罚都与投票人的投票方式无关。

众所周知，一个人应该选择一位谈判高手为自己服务，并授之明确的权限。但是这条原则通常出自谈判人员口中，当然这也绝不可能如谈判对手建议的那样不言而喻。谈判人员的权限大多在谈判中无法做出让步，以及在提出要求时显得捉襟见肘。[1]同样，尽管审慎告诉我们，在以自杀式报复作为威胁的手段时，最好给自己留条后路以防万一，但是任何留有后路的迹象都是在削弱威胁的可信度。在众人看来，主动放弃某种选择权，甚至放弃对未来命运的全部控制权并自动做出反应而具有战略优势似乎很难让人理解。

现实中，许多案例涉及对谈判技巧、智慧、理性、知识、控制力甚至自由的否定。原则上讲，这些在某种条件下都是有效的。而且如果将问题公式化，对此进行抽象的研究，并把其中的类似性拿到其他陌生感没有成为理解障碍的语境中进行检验，那么，克服陌生感、理解其背后的合理性就变得更加容易了。

另一个与第一印象截然不同的原则涉及干净核弹和脏弹的相对优势问题。伯纳德·布罗迪（Bernard Brodie）指出，如果考虑到威慑与所希望进行的战争之间不同的特殊要求，则可以考虑使用超脏核弹。正如第10章中所言，如果我们把“恐怖平衡”（balance of terror）简单地看做是古代人质交换机制的大型现代版，这一结论也就理所当然了。

[1] 在对外援助管理中存在大量这方面的案例。例如.C.Schelling,“American Foreign Assistance”,*World Politics* (July 1955) pp.614~615。

这里，我们也许发现了当今文明时代国际政治专业学生的一个弱点。当然，这个弱点是相对于古代的马基雅维利和中国人而言。我们需要通过一些诸如信任、善意和相互尊重等概念区别和平、稳定和冲突的平息。可以在某种程度上讲，这一观点有助于推动互信和相互尊重，因此这是件好事。但是在互信和善意根本不存在时，或者我们虽努力却无法实现互信和善意时，我们期望能从黑社会或古代的独裁者那里得到启发，学会在互信和善意无法存在和缺乏维护契约的法律保证时，如何才能达成妥协。古人的很多做法值得我们学习，如交换人质以示诚意、同饮一杯水以示清白、到公共场所会面以防被对方暗算，甚至人为地交换间谍以方便真实情报的传递等。完整、系统的战略理论有助于现代人认识和了解古人的做法，揭示其有效条件并研究发现当代的有效条件。尽管这不是我们的初衷，但或许正是冲突管理核心之所在。

第2章

论　谈　判

本章主要讨论有关谈判过程中的战术技巧问题。谈判既包括谈判主体之间的唇枪舌剑，也包括二者之间无声无息的心理斗争。后者中，谈判主体都明白，自己的一举一动都将对对方产生影响，从而试图通过各种具体行为来达到自己预期的目标。经济学意义上的谈判包括劳资谈判、关税谈判、少数竞争者间的庭外调解、房产商与顾客之间的谈判等。经济学之外的谈判含义十分广泛，从大规模报复行动到出租车司机的通行权等问题。

我们这里所指的谈判不是调整谈判主体间经济利益分配的谈判。后者涉及谈判的功效性（efficiency）问题。例如，保险公司是否可以通过向投保人提供现金，而非修理被损坏的汽车来达到既节省成本又令投保人高兴的目的？雇主是否可以通过向愿意以实物顶替薪水的雇员提前增加薪水来达到节省成本的目的？相反，我们关注的是谈判的分配（*distributional*）效用问题，即：在什么条件下，谈判对主体一方更有利。当交易完成，商品从商家手中转移到购买者手中后，价格还有什么意义呢？当两辆卡车相向而驶，但是道路仅可通过一辆卡车时，哪辆卡车会让路呢？

上述情景最终都涉及到完全谈判。在完全谈判中，谈判主体双方很大程度上都会受到自己对彼此行为预期判断的影响。然而，由于谈判主

体每一方都受到自我对对方行为预期判断的影响，并且知道对方亦然，因此双方对彼此行为的预期就会发生交叉影响。当一方最后一次做出充分的让步时，一场谈判也就宣告结束。那么，他为什么要首先让步呢？原因在于他认为对方将不会让步。“我必须让步，因为对方不让步。对方不让步，因为他认为我一定会让步。他认为我会让步，因为他认为我认为他是这么认为的。”看起来双方之间似乎存在很大的选择余地，并会产生不同结果。但是无论产生何种结果，总比双方之间没有达成任何共识要好。这是“完全谈判”成立的核心条件，因为总有一方愿意率先妥协让步，双方从而达成某种共识。如果双方都知道达成妥协的临界点，那么任何一种结果都是至少一方愿意妥协接受的结果，对此对方也会心领神会。如果这样的话，也就不存在任何谈判僵局的问题。

因此，实际谈判一定能够达成某种结果。这种情况下，我们也许在逻辑上找不到答案，但是我们可能在运用的策略上找到想要的答案。本章的目的也是为了介绍一组重要的谈判策略，而且特别适用于不确定的情况。本质上讲，这组策略是建立在谈判主体自愿原则的基础上，而且要求主体牺牲自我的自由选择权。这意味着遏制对手的权力就是限制自己的权力。在谈判中，优势也是弱势，自由也许只是让步的自由，破釜沉舟也许意味着给对方更多的机会。

谈判实力：自我约束的实力

“谈判能力”、“谈判优势”、“谈判技巧”都似乎暗含，谈判的优势属于拥有实力、优势和技巧的谈判主体。如果把谈判的胜利局限于获胜的一方，上述观点无疑是正确的。但是，如果认为谈判中拥有更多智慧和技巧的一方，或谈判中拥有更强大的经济、物质基础、军事力量以及打击承受能力的一方占有优势，上述的观点就失去了积极意义。因此，对谈判主体而言，这些条件绝不是在任何谈判中都是有利的，有些时候往往适得其反。

一个经验丰富的谈判家也许会发现，要成为一名真正立场坚定的判断者十分困难。如果有人上门威胁主人说，要是不给 10 美元，就会捅他一刀。主人如果看到来人满脸杀气腾腾，也许真会妥协。鉴于此，当

对手过于狡猾，能够识别威慑时；或者对方过于软弱，根本无力保护自我时，那就完全没有必要使用威慑。一个无力保证收支平衡、税收、运用暴力机器保卫自己国家的政府也许会得到国际社会的援助，但是援助国如果无法控制该国的资源，也会拒绝提供援助。这里引用一个大家比较熟悉的经济学术语“价格领先”（price leadership）。在供不应求的市场中，“价格领先”意味着小企业无力可图而逃之夭夭和大企业的“欺行霸市”。

讨价谈判的实力也可被看做是欺骗（bluff）、愚弄（fool）的实力，“这种能力也指自己制定最有利的价格，并骗取对方相信这是自己所能承受的最终价格”。[1]虽然欺骗和愚弄不离其中，但是欺骗分为两种：一种是隐瞒事实。如顾客故意隐瞒自己的收入和家庭成员；另一种是策略性的。假设商家和顾客对彼此十分了解，一方如何欺骗另一方呢？尽管买方能够承担 20 美元并知道卖方也知道自己的情况，但是他还是可能会说自己最多能够出 16 美元。如果卖方让步，难道我们可以说卖方被欺骗了？难道卖方不知道真实情况吗？如果买方的骗术被卖方识破，难道买方不知道下一步怎么做吗？假设买方感到（feels）自己非常坚决，并以卖方一定会让步作为其理由。一旦卖方妥协让步，事后买方也许会认为卖方并没有被“愚弄”。无论怎么说，欺骗和愚弄观念在此都不完全适用。

如何才能让一个人相信一件事呢？答案很大程度上取决于现实条件。“正确吗？”事实表明，向一个人证明一件事正确比向其证明一件事错误要容易得多。要证明自己健康，我们只需要找来一位合格的医生作证。要证明自己的成本和收入情况，我们只需要让一个人查查经过国家税务局或一个信誉良好的会计公司审计过的账本。但是，要说服一个人相信一件事是错误的，我们可能很难找到令人信服的证据。当一个人试图说服另一个人，他只能支付 16 000 美元去买一栋价值 2 万美元的房子时，他该如何利用真话比假话更令人可信的优势呢？答案是：努力

[1] (J.N.Morgan) ,Bilateral Monopoly and the Competitive Output,*Quarterly Journal of Economics*; 63:376n6 (August 1949) .

让它听起来更像是真话。那么，买方又是如何做到这一点的呢？如果房子的位置靠近买方的工作场所，他会避口不谈自己的工作，并努力说服卖方，房子对他来说现在只值 16 000 美元，这已是自己能够承担的最高价钱了。如果他付出高价买了这栋房子，还不如到别处另寻更好的居所呢。

现在我们假设：买方已经与第三方达成了不能反悔的协议，并且已经生效。根据双方协议，买方的支付不会超过 16 000 美元，否则买方将输掉 5000 美元。如果这样的话，卖方就输了，因为买方需要做的只是将事实告诉卖方而已。除非卖方懊恼逞凶、恶意占有房子不愿出让，否则形势对他十分不利。“客观”形势——买方的真实意图——已经主动、明确、完全地发生逆转。卖方要么继续保留房子，要么低价卖掉。这个案例揭示，如果买方能接受一个不可变更的契约，卖方也知道这一事实，买方就可以把谈判的空间缩小到对自己最有利的状态。此外，这个案例还说明，由于策略本身的欺骗性，策略可能存在，也可能不存在。买方能否找到一个能够使自己履行契约的有效方式，很大程度上取决于买卖双方的身份、居所，以及众多法律保障的机制安排（包括在本假设案例中，买方与第三者达成的协议是否具有法律效力）。

如果谈判主体共同生活在一个“真话”（cross my heart）被普遍认为具有法律效力的文化氛围内，买方需要做的是坚持自己的支付能力不能超过 16 000 美元，并发誓如有隐瞒，将受到严厉的惩罚。这样，买方就赢了，至少卖方不会用大喊大叫“19 000美元，我发誓”来要挟买方。假设买方是受董事会委托只能以 16 000 美元价格购买该房的代理人，而且董事们根据公司章程又不能在几个月内再次开会，买方也不能私自越权加价，同时，卖方也熟悉这一情况，那么，只要出现下列情况之一，买方就是“赢家”：卖方没有将自己的要价锁死在 19 000 美元；或者买方坚持自己只能支付 16 000美元，并且不愿因多支付一美分而失去个人良好声望和谈判信誉，因为他多支付的事实可能会被别人得知或卖主一时高兴把消息捅了出去，以迫使买方履行承诺。当然，除非卖方对情况了解得十分清楚，否则这种方式无须牺牲灵活性。

契约式承诺有时并非如第一印象那样有效。在前面那个带有自罚性质的案例中，卖方完全可以找出第三者，提出一个适中的报价，并威性

他：如果不能免除买方的违约责任，他将以 16 000 美元出售房子。如此，买方与第三方所达成的协议——与大多数契约式承诺一样——转换了谈判中涉及的房子与主体的关系，旨在避免第三方涉入谈判或做出让步。换句话说，契约式承诺通常涉及假设的暂时“转移成本”（transfer cost）问题，但并非“真实成本”（real cost）以及当事方是否愿意卷入无休止的谈判之中。假如第三方以巨大的“转移成本”出现，承诺通常也会被履行。（假设该协议涉及多人，鉴于高昂的“真实成本”，他们也许不愿意卷入谈判。）[1]

我们的话题中最有趣的一部分，涉及是否履行以及如何履行承诺。首先，我们有必要假设一个没有现实缺陷的完美状态。在这个完美状态中，完全承诺可以自由存在。我们还得考虑一种“说真话”被普遍接受具有绝对效力的文化环境。在这种文化环境中，谈判一方提出的报价都是最终报价，并将被另一方接受。如果一方知道另一方的真实价格，目标就是第一报价必须是实盘（firm offer)，后果的全部责任则属于接受或拒绝报价的另一方。这时，整个谈判过程宣告结束，最初的承诺方取胜。

在此，我们有必要提出沟通方面的一些注意事项。谈判双方一定要通过信件进行沟通。虽然当一方签字时，承诺就已经生效，但是另一方这时并不知道，因为他还没有收到信件。现在的情况是，一方还在起草

[1] 也许解决双边垄断问题的最佳方式是，一方转移其边际成本曲线，从而使共同利润回到其最大值的最初起点零。一方可能通过不可更改的售后回租合同实现这一目的，即向第三方出卖特许权合同以套取大量现金，而特许权与其产出相关联，共同成本大于其他产出的共同税费。现在，他在任何价格下都无法生产，除非这一价格或生产能够保证最初的共同利润属于其本人。另一方看到这一合同，也会表示同意，并接受其真实的最小利润。真正的“赢者”通过出卖特许权而获得全部原始利润，利润并不影响他的动机，因为这与他的生产无关。第三方支付大量现金（除去回扣），因为他知道第二方一定会让步，自己会最终赢得暂时的特许权费。问题在于，获得特许权的一方并非“长期一员”，否则后者会威胁与他终止特许关系，并拒绝任何协调，回到最初的边际成本状况。但是，我们可以假设机制的完善能够使特许专业化，从而最终的胜利依赖于从不食言、具有良好声誉的一方，并且不被卷入任何谈判之中。

信稿时或第一封信到达之前，另一方或许已经提出了条件，结果，却被告知没有现货。可以说，双方处于信息极不对称的位置。任何一方都必须做好心理准备，准备面对僵局或对方已经另有意向的局面。

尽管如此，信息的不对称性或许对尚未收到信函的一方十分有利，因为他不需要承担接受对方提出的条件后又反悔的责任。（换句话说，如果一方无法进行有效沟通，就可以为自己的不作为开脱责任。另一方也可以以自己对对方条件不详为由，而拒绝提出条件。）如果承诺不仅可以体现在语言上，也可体现在某种形式或行为上，那么在一方的不作为能被充分理解的情况下，其对承诺的不作为绝对是一个有利条件，因为另一方明白只有自己做出让步才能避免僵局出现。

假设整个国民中，只有一部分人群“说真话”具有（或被认为具有）效力，那么，其中对他人情况了如指掌（或应该知道）的那部分人就占有绝对的优势，因为他们可以轻松地做出决定，而其他人却不能。如果他说“16 000 美元，我向上帝发誓，这绝对是真实的价格”，他的报价也就是最终价格。如果他说“19 000 美元”，那他只是（或应该是）在“讨价还价”。

在普通的讨价还价过程中，如果一方不知道另一方的真实价格，双方之间也许存在一个互相探询对方底线的过程。但是，这一过程会很快发展为相互发盘与还盘的过程。由于各种实际原因，双方的讨价还价总会影响“真实”的价格底线。假设一方相信某种履行承诺的方式，而另一方则没有，那么后者往往追求诸如“公开”其底线价格等“一般”谈判技巧，而前者将继续坚持自己的立场。

前面的探讨旨在向读者揭示自我承诺（self-commitment）的可行性和合理性。尽管许多案例涉及了相关的谈判策略与技巧，但是一些研究人员很少能自信地区分相关策略的合理（logical）运用、本能（intuitive）运用和无意识（inadvertent）运用之间的异同之处。例如，工会领导者通常不会在劳资谈判期间或之前为工会成员的利益而兴奋不已、坚定立场。如果工会希望得到2美元，并期望管理部门还价是1.6美元，那么工会有必要说服工人相信，管理部门能够支持2美元，一旦无法达到这个目标，那就是谈判人员的无能。工会这么做的目的——我们研究得到

的可行性结论——也是为了让管理部门明白，即使谈判人员能够接受低于2美元的条件，他们也只能坚持2美元的立场，因为他们无法控制工人的决定；或者如果让步，他们就难咎其职了。换句话说，工会谈判人员人为地缩小了自己的权限，并以工会无法阻止罢工威胁管理部门，看起来这似乎是工会主动放弃阻止罢工的权力。

上述的情景似乎也存在于美国与其他国家的政府间谈判，如前者向后者提供援助或削弱关税谈判。如果行政部门拥有充分的权限决定自己的谈判活动，谈判就不会导致任何所谓的罢工活动，而以存有争议的妥协立场结束。因为谈判对手知道，美国政府宁愿妥协结束谈判也不愿无疾而终。但是，如果行政部门在司法部门的监督下进行自己的谈判活动，并将受到法律的约束，而国会不可能在必要的谈判期间修订法律来满足行政部门谈判的需要，那么，行政部门的强硬立场就会让谈判对手清晰可见。

代表国家利益参加各种国际谈判的谈判代表通常都知道，国际谈判存在着很大的谈判空间，谈判主体的讨价还价也会产生不同的结果。在谈判正式开始之前，谈判人员总是设法发表各种声明制造气氛，强调自己的坚定立场，从而激起民意，为自己在谈判中创造有利条件。一旦这种造势活动成功并将某种信息传达给谈判对手，那么，最初的谈判立场似乎看起来更像是终极结果。

这些案例都呈现一些相同的特点。第一，谈判一方并不急于得到最终的结果，他们更重视与谈判对手的沟通。第二，做出最终决定对双方而言显然是一件非常困难的事情，更不用说谈判主体了解对方的真实意图和立场。第三，谈判正式开始之前，谈判双方可能都在进行同样的准备活动。第四，尽管谈判双方都有提出条件的机会，但是机会对双方而言并非均等。由于民主政府受到民意的限制，提出条件的能力显然无法与集权政府相比。最后，由于谈判双方提出的强硬条件有可能超过对方的接受能力，因而双方都存在可能导致谈判出现僵局或谈判失败的风险。

谈判机制及其结构特点

谈判本身特有的一些机制和结构特点可能造成提出条件的策略和技

巧对一方有利，而对另一方相对不利的局面，甚至造成谈判双方同时提出条件或谈判出现僵局。

巧用谈判代表　谈判代表至少以两种方式影响双方提出条件的权威性。一方面，谈判代表通常会接受难以更改或根本无法更改的指示，而且这些指示（或者说谈判人员的非灵活性）易于暴露给谈判对手。这一原则适用于行政部门与司法部门或管理层与董事会之间，特别是使节出席的谈判场合，即在时间紧迫的谈判过程中，虽然主要谈判人员与使节之间存在较大分歧，但在谈判结束之前，双方无法进行有效沟通。

另一方面，接受委托的“代理人”必须能够在其权限内自由决定谈判事项，即使其与被代理人在谈判期间就谈判的原则立场存在分歧。汽车保险问题就属于这一类情况。参加庭外和解的投保人如果以上诉威胁保险公司，其威胁显然无法对后者产生很大的效果，因为为了维护保险公司以后的声誉和形象，后者会不惜一切代价进行还击。[1]

保密与公开　提出条件的一个有效方式——有时可能是惟一方式——是假借维护谈判主体的声誉和尊严。如果代表国家利益的谈判代表有权做出较小的让步并与其个人奖惩密切联系，他们就会千方百计将让步控制在自己允许的范围内。如果工会在与企业主的谈判中做出极大的让步，其威望和声誉将受到严重的挑战，并失去进一步妥协的余地。（同样的道理适应于普遍公开化的国防事务。“如果你这么做了，我就必须以同样的方式对待别人。”）但是，有时谈判也需要适度地对外公开，如谈判的最初条件和最终结果。一旦保密事项涉及谈判核心问题或谈判结果一时无法预见，这个方法就失去了作用。如果谈判一方具备对外公开的条件，而另一方不具备这样的条件，后者就会竭尽全力避免相关的对外公开事宜以扭转自己的不利处境。如果双方都担心彼此同时使用这种方法将造成谈判出现僵局，那么他们会努力就相关的保密事项达

[1] 交通案例中的通行权问题的解决方法是，胜者是第一个为汽车投全险的司机，因为他不必惧怕任何交通事故，而另一方必须妥协、退让。（后者没有还价的余地，因为前者已经投了全险，现在没有保险公司愿意为其投保。）更严重的情况是，工会罢工基金的减少，降低了每个工会阻止罢工的动力。双边垄断的解决方式反映了对第三者存在转移利益的问题，并影响一方动机结构的变化。

成某种共识。

交叉谈判　如果工会有能力（即使是暂时性地）同时参与几个谈判，并敢于以自己的谈判声誉冒险，而管理部门却不具备这样的条件，那么，率先主动转移到其他谈判问题的一方就掌握了谈判的主动权，即使该方在这些谈判中做出妥协让步，其强硬立场也将超过对方的承受能力。［与谈判双方在谈判最初阶段重视的立场坚定性相比，谈判主体的“名誉价值”（reputation value）与谈判结果的关系相对松散。］双方之间策略与反策略或技巧与反技巧并存，例如谈判双方都有可能人为地对对方立场产生误解，或努力达成一个与最初的立场不同的结果。如果谈判中，谈判的内容被扩大化，例如劳资谈判成了额外福利谈判，这就背离了原来的谈判初衷，那么只好请首先提出另外议题的一方“出局”(out)，而“出局”则削弱了该方的谈判地位并对其产生不利影响。

持续谈判　国际谈判中有时存在这么一个特殊场景：谈判双方在就一个问题进行谈判的同时或者在不久的将来，还就多个问题进行谈判。这一谈判场景的逻辑结构十分微妙。为了说服谈判对手相信自己已经退到了谈判的底线，一方可以巧妙地说：“如果我方在这个问题上做出让步，你方就会在其他问题上得寸进尺。为了尊严和荣誉，我方必须坚定自己的立场。”这样的话，谈判对手现在同时成了维护其谈判尊严的“第三方”。当然，通常只有在地区性谈判中才会出现这一情况。实施威胁的谈判一方不仅达到了自己的目的，并且就提高未来威慑效果的长期性（而非暂时得到的利益）而言，也加强了威慑的效果。

限制性议程　当谈判涉及两个主题时，双方是否决定同时，或同时不同地，或在不同时间进行相关问题的谈判与谈判结果没有丝毫联系，特别是在合法的普通谈判中，某种潜在敲竹杠式的威胁存在时。对抗威胁的方式多种多样，包括拒绝对方的条件、谈判中不作为，甚至终止谈判。但是，如果威胁的对象涉及其他谈判问题的议程，潜在威胁就会变得更加有效。

关税谈判就是一个典型案例。如果谈判的焦点是干酪和汽车关税，谈判一方也许会以提高其他商品的关税威胁对方以达成对自方有利的结果。但是如果被威胁一方仅仅被授权进行前两个问题的谈判；或者在谈

判之前双方已经达成相关协议，即在双方尚未就前面两个问题达成共识前，不得进行其他问题的谈判，那么，威胁主体的伎俩只好另待良机。如果威胁的方式不容许对外公开，那么，来自公开化的压力也会妨碍其发挥应有的作用。

补偿的可能性　费勒（Fellner）曾经指出，谈判主体能否达成共识很大程度上取决于成本和利益的再分配方式。[1]例如，在双边垄断的市场环境中，双边寡头分割市场的原则是实现共同利益的最大化，其中最初阶段的利润增长点最为关键；其他市场划分方式要求其中一个公司必须给予另一方适当的经济补偿。但是，一旦补偿行为成为非法共谋的证据，或补偿的动机被股东误解，或双方对彼此存有疑心，那么，双方对公平占有最初利润增长点的要求势必会对双方共有利益产生不利影响。

如果一定要分配属于某人的东西时，则分配的成本取决于补偿的程度。在这种情况下，谈判的“议程”（agenda）就变得尤其重要，因为补偿的主要方式是一方在其他问题上的妥协让步。如果两个同时进行的谈判密切了双方的关系，二者就有可能达成某种补偿结果；反之，每一方只能继续独占自己的所属物。

在某种情况下，也许就不同的问题分别进行谈判对谈判一方有利，而另一方则希望同时进行两个问题的谈判。例如，现在有两份都价值 3 美元的物品需要分配。A 得到的部分价值 2 美元，而 B 得到的部分价值 4 美元；在具体分配过程中，分配后的物品成为持有者的“个体”(one-man）财产。如果补偿在机制上无法实现，只要两件物品被分别保管，那么 B 将被迫承担全部分配成本。B 无法以无作为威胁 A，因为哪部分 A 都不想单独占有。但是如果 B 将双方的占有物放在一起，提出自己占有一部分，而 A 占有剩余的部分，并有效地威胁 A 必须其中二选一，否则自己也不会占有。现在 A 面临惟一的选择，占有价值 4 美元的那部分并同时承担 3 美元的分配成本。无论 A 占有哪部分，B 都将自己分摊的分配成本减少一半。

作为谈判的典型，经济学问题也存在一个重要缺陷，它们总是试图

[1]（W.Feller）,*Competition Among the Few* (New York，1949)，pp34~35，191~197，231~232，234.

将不可分配的东西与补偿行为联系起来。例如，在一家房屋后面挖一条排水沟对其附近的两家都有好处。一家需要支付 1 000 美元，另一家则需要支付 800 美元，但是两家谁也不愿支付自己的那部分。现在我们不妨假设两家走到一起，把这一工程看做一个价值 1 600 美元需要双方共同努力的项目。又如，童子军团长（scoutmaster）一职需要占用一个人一周的 10 个小时，而每个人都认为自己只需要 8 个小时参加童子团，但是必须有一个人负责全部的工作，那么，我们很难想像两个邻居之间会达成这样一个协议：一方花费 10 个小时用于工作，而另一个人愿意为其支付现金或为其放哨 5 个小时。两辆汽车狭路相逢时，双方都缺乏主动让路的精神而加剧堵车。当议会中互投赞成票不适用时，往往也会出现僵局。要求一致同意的措施往往只有几个人团结起来才有可能被实施。[1]

谈判技巧　尽管谈判具有许多值得研究的特点，但是我们尚无法完全发掘。传递错误的信息应该受到惩罚吗？假设发盘一方在对方接受后反悔是欺骗行为，那么，是否应该惩罚这种欺骗行为？是否应该惩罚雇用他人充当“托儿”来探寻对方情况的一方？是否能够承认所有相关的当事人？谈判应该有限制时间吗？谈判应该采取哪种方式，拍卖式、自动落价拍卖式、封闭的招标式，还是其他形式？缺席谈判的一方是否能够赢得优势？谈判出现僵局时，双方是否可以恢复谈判？谈判出现僵局的后果是什么？能否监督双方对协议的实施？一般而言，谈判双方运用何种沟通方式可能造成信息失误？如果谈判双方存在多个谈判问题，他们能否同时进行相关谈判，能否分开逐个进行谈判，或能否同时不同地以不同规则进行谈判？

如果我们认真研究议员们常用的政治技巧和策略，上述结构性问题的重要性就不言自明。诸如总统只能完全否决一个拨款议案，或要求在整个草案表决之前，必须对修正条款进行逐个表决，以及优先制度必须根据不同的动因实施等规定都极大地影响了每项动案背后的动机。有人

[1] 相关的内容有“巴黎和约”中有关萨尔地区的条款。该条款规定结束德国对萨尔地区的占领。完全符合此处以及前文所提到的这一原则。

可能被迫选择第二种方式，但鉴于这是他为了消除某种可能性而自愿提前采取的表决行为，多少会令其感到欣慰。但是，他更倾向于剩下的第一和第三选择项，所以，不会产生任何外来威胁。

原则和先例　为了增强立场的可信度，通常要求谈判主体提出的条件高质量而非多数量，并且应该具有一定合理性。例如，很难想像一家公司的报价是 2.07 1/2 美元，为什么不是 2.02 1/4 美元呢？这样的数字很容易让人产生怀疑，除非是像 2.00 美元这样的数字。尽管如此，诸如“分红”（profit sharing）、“生活成本增加”（cost-of-living increases）等理由也许能使人相信可以得到 2.07 1/2 这样的数字。此外，谈判主体也可能提出将一些有利原则和先例置于危险境地的条件。过去，一方可能成功运用了某些原则，诸如否认武力夺取政权的合法性。现在，他还想继续将此原则运用于谈判之中，但适得其反，不仅削弱了自己的立场，而且还使对方开始怀疑这一原则的有效性。但是为了维护其有效性，他仍可能坚持说服对方，宁愿谈判出现僵局也不妥协。

诡辩　如果谈判一方认为谈判到了必须妥协的地步，那他必须牢记妥协造成的两个后果：1.妥协将促使自己的立场更接近对方的立场。2.妥协将导致对方怀疑自己立场的坚定性。让步不仅被对方释义为软弱，还将导致对方怀疑此前所有条件的真实性，而且进一步怀疑新条件的真实性。因此，谈判主体应该寻找一个“合理”的理由来说服对方，最好能对自己此前的立场进行理性化重释。

更有趣的是，谈判主体一方可能有意（或无意）运用诡辩为对方开脱责任。如果一方向对方表明后者可以不履行承诺或自己提出的条件值得商榷，某种程度上，这是他在帮助对方开脱责任。当然，谈判一方也有可能混淆对方的立场条件，从而导致自己的委托人或谈判代表或公众，也无法准确理解某些立场条件——“生产力”（productivity）也许太含糊不清，而“按比例捐献”（proportionate contributions）又存在歧义——总之，他在诋毁或削弱承诺的有效性。当然，在这种情况下，如果有人成功驳倒这些结论，也会令谈判另一方处于不利境地。但是，当对方决定做出适度让步时，另一方可以通过某种方式协助对方肯定其适度妥协的“必要性”，从而促使谈判达成对自己有利的结果。如果即便

为对方开脱责任也无法说服人们相信对方妥协的合理性，那么，对方将继续坚持以前的立场。换句话说，谈判一方必须学会寻求一种合理化解释才能说服别人，否则将打击对方妥协让步的积极性，即便对方勉强做出某些让步，自己也不可能从中获益。[1]

威　胁

人们受到对手的打击威胁或受到竞争者的降价威胁时，通常都会奋不顾身地进行报复，这等于主动暴露自己的真实意图，迫使对方自食其恶果。但是更奇妙的是，这种威慑行为一旦成功，往往对双方都有利。

如果一方只进行口头上的自杀式报复威胁而没有实际行动，这就不仅仅是沟通问题了。例如，一方也许会以大规模行动报复对方微弱的侵犯举动。类似的情况还有撞坏汽车也得不到通行权；即使举行规模宏大的罢工活动，也不一定增加几美分薪水。这些威胁行为最显著的特征是：威胁主体从未有付诸行动的决心和意志。当然，如果宣传上的威胁就能达到预期目的，也就不存在实施的必要性了。威胁偶然实施的可能性越大，真正实施的可能性就越小。但是，一旦威胁的效果取决于对方是否容易轻信，那么威胁方必须重新组合或调整自己的威慑动机，向对方显示自己实施威胁的决心。否则，威胁也会丧失其自身价

[1] 许多课本中提到了相关问题，如公司双边垄断问题，认为谈判是一种零和博弈，一方之所失即另一方之所得，强调一方微乎其微的作用和立场不如根本不强调。但是，除了某种消费和销售情况外，很多限制条件会影响双方可以接受的谈判结果，即使达成对一方最不利但却可以接受的谈判结果，也比无果而终要好得多。在这种情况下，一方的首要目标是预防对方的阴谋诡计。如果真象比假象更容易识别，则有必要在一开始就采取比较保守的立场，否则一旦让步就会对以后的立场产生不利影响。事实上，尽管没有人愿意为自己画地为牢，但如能巧妙地向对方展示自我惩罚的姿态，往往会产生积极的效果。如果一方试图向别人证明自己的成本或收益情况，可能会通过展示自己的个人所得税情况，而且这时如能适当地出现一场自我惩罚的假戏的话，效果会更好。

甚至在某种程度上，假设谈判人员是其委托人或公司职员，而且其收益标准是达成协议，而非达成协议的效用性，那么，“完全”双边垄断问题就更符合这一标准。

值。[1]

现在，让我们再回到最初的承诺问题。谁愿意为了威慑对手，把自己过早地束缚于一个自己毫无实施意愿的威胁承诺？当然，一方可以以欺骗手段说服别人相信，自己的威胁行为对对方的消极影响或损失微小或根本就没有，而这样就降低了威胁的必要性。更有趣的是，威胁主体可能会假装自己之所以会错误地做出威慑承诺，完全是因为自己错误估计了实施威胁的后果；或许，他应该着重强调实施威胁是为了保护自己。但是在别人看来，这些理由可能已经成为报复行动的真实依据，那么，他必须再为自己的行为寻找其他充分理由。

一方或许敢于冒险做出威胁承诺，即使冒着丧失尊严和荣誉的风险，也要令对方受到震慑。如果没有认识到实施威胁所付出的代价，他甚至会采取两败俱伤的威胁行为，只要他认为这一切能给对方以沉重教训。此外，他还可以通过合法途径将实施威胁的任务转移给实施威胁的“代理商”来完成。[2]尽管后者会为此付出一定的代价，但是前者最终还是要承担主要的责任。

法律术语“最后的机会”（last clear chance）或许能够充分说明承诺问题。“最后的机会”是指在交通事故发生之前，对肇事双方而言，存在一个可能避免事故发生的临界点，这是他们阻止事故发生的最后机会。谈判中，一方过早地做出最后承诺也许是放弃了与对方达成共识的最后机会。这也意味着他放弃了任何妥协、让步的机会，将过早地僵化自己的立场，但是对方也可能必须做出对其有利的选择。例如，一个司机已经加速行驶并且无法刹车，另一位司机认识到了这一点，也只好妥协让步。立法议员也许在会议最后阶段给总统一次最后悔过机会批准某个法案，而后者只能接受，因为无论其接受与否，该法案都将生效。这

[1] 偶然的是，威胁行为具有很多有趣的特点，这些特点反映了回报与惩罚之间的不对称性。例如，威胁方的收益也许小于被威胁方。这如同以一辆新车撞一辆旧车，或为了微小的损失而兴师动众地进行价格战。因此，如果威胁的成本过高，就没有实施的必要性。当然，以宰牛刀杀鸡也许有些荒谬，但具有合理性，只要被威胁方认为威胁难以抵抗而主动放弃。

[2] 如果以这种方式理解，弱国与强国之间的防务条约并不是强国为了保证弱国的安全或掠夺弱国的补偿，而是弱国牺牲自由权的一种妥协方式。

一原则有助于我们理解在一些情况下，为什么一定意义上的弱势（优势）能够转化为另一个意义上的优势（弱势）。当一个人或国家失去自助能力或失去避免与敌人同归于尽的能力时，另一方必须为此承担全部成本和责任。“强制不足”（oercive deficiency）是奥瑟·史密西斯(Authur Smithies) 用来形容一个国家为了造成资金紧缺的局面，而人为地将年度预算提前消耗掉的一种策略。[1]

类似的策略还有如何将谈判引入到只有某种过分行为才能迫使一方让步的环境（status quo)，因为谈判双方一旦进入这一阶段就没有任何后路可言，任何过分的行为都只会加快两败俱伤的情况发生。如果一个人身上带有易燃炸药，并且不可避免地造成自己与任何侵犯者同归于尽，那么只要存在一线控制身上炸药的希望，他就只会以同归于尽相威胁，而不会引爆身上的炸药。同样，即使一方派出的只是象征性的武装力量，但只要存在无法收回的可能性，也会激起对方的抵触情绪。沃尔特·李普曼（Walter Lippmann）曾把它比作保护珠宝店的平板玻璃：只要任何人都可能打碎它，就无论如何避免不了引起混乱。

被威胁方同样也可能使用类似的策略和技巧化解威胁。被威胁方的防务力量当然是在威胁到来之前作好准备，因为只有这样，才能化解威胁方报复的意图和意志。如果他无法实现最短时间内的防御准备，就只能奋力一搏。当被威胁方决定孤注一掷时，威胁方也就不能放弃威慑承诺而只能继续行动，即便两败俱伤。[2]一旦被威胁方也可在威胁到来之前，将威胁的风险转嫁给第三方（如前面提到的以投保方式解决通路权

[1] A.Smithies,*The Budgetary Process in the United States* (New York,1955) ,pp.40,56.他认为，一种有效方法是缩短分配的过程。请参阅 T.C.Schelling, “American Foreign Assistance,” *World Politics*,7:607~625 (July 1955)，同样涉及对外援助的分配原则问题。

[2] 对管理官员来说，通过在与司机交涉之前或在其威胁发生之前记下他的车牌号，为警察提供一定量的通行权和免责权的制度是可行的。有些卡车写着“司机无法控制警报系统”，这与银行里的定时锁以及强制实施无记名选举具有同样的效果，这如同以小股军事力量进行侵犯行为，即使军事力量规模小到无法达到预期目标。企业也是一样，妥协往往与“脸面”有很大联系：强势一方往往不会担心任何威胁。在很多大学里，为了保护职工的利益，禁止导师修改已经记录在案的分数。

问题），即可降低自己受威胁的程度，迫使威胁方转移威胁的对象；或者以其他方式重释自己的立场，并对外造成一种错觉：即使对方真正实施威胁，自己也会取得最终的胜利（或许他自己这么认为）。鉴于威胁的高昂代价以及无果而终的可能性，威胁方也许会知难而退。同样，如果一方无法代表自己的利益，或无法正确解读威胁方的真实意图或动机，或坚持对对方视而不见，这也可能避免对方的威胁。当然，最好的对策是表现出“真正”(genuine)无知、固执己见或轻信，因为这样更容易迷惑潜在威胁方。这个方法只能成功，不能失败，因为一旦失败，双方俱损。最后，威胁和威胁承诺必须让双方知道。要是被威胁方信息渠道不畅通，或故意破坏沟通渠道，正如其出发点是为了避免威胁，他确实能阻止威胁产生。[1]被威胁方必须时刻牢记：向对方展示自己的坚定立场或对对方的蔑视必须在威胁到来之前；否则，一切都于事无补。

在许多威胁情况下，如常见的谈判、威胁承诺含糊不清；双方往往都不能准确地判断威胁给双方造成的损失和影响；如果威胁承诺是一个持续的过程，那么承诺的坚定性必须体现在一系列行动上。这时，双方的沟通既不是完全没有可能，也不是完全可靠，因为一方的某些意图能够被直接传达给对方，有些则只能通过报纸或其他间接方式，或者直接通过行为本身传递。一旦出现这种情况，双方都极有可能做出令双方不开心的举动，如同时实施威胁与反威胁。而且，一旦双方都认识到了这一点，就会设法阻止双方同时采取实际行动。[2]

一旦威胁承诺已经做出且无法撤回，威慑主体在威胁行为付诸实施之前，还存在最后一个避免两败俱伤的机会，即不作为(undoing)。这

[1] 敲诈者也许因为主人不在家，而失去敲诈的机会。绑架犯也是同样的情况，如果无法与被绑架者的家人或朋友联系，也达不到目的。因此，法律禁止被绑架者家人或朋友与绑架者联系，以达到打击绑架者的规定也许没有实际意义。同样的道理，监督人员或警备人员轮流值班，或随机组合值班人员不仅能够防止他们受贿犯罪，也能防止他们受到威胁。

[2] 就这里我们所讨论的承诺而言，在人与人之间或国与国之间实施方式多种多样，并不惟一。但是，大多数情况是非常模糊的，偶而能够发挥作用。在前面提到的崇尚“真话”的社会里，谈判理论仅仅属于博弈战略和沟通机制。然而，在当今世界，谈判主要指拥有实力保证并被对方所认可的经验性的机制过程。

时，威胁的初衷已经消失，威慑的收益也转变为零，只有此前做出的威胁承诺尚在，并成为实施威胁的惟一动力。这类似于在普通谈判中，因谈判双方无法协调立场或一方提出超过对方承受能力的条件而出现的僵局。如果存在不作为的可能性，双方都会努力寻求共识。但是，不作为往往涉及双方之间利益的分配，因为不同方式的不作为将导致不同的结果。而且，“不作为”并不意味着抛弃尊严和荣誉而拒绝履行承诺；相反，如果实际存在名誉承诺（the commitment of reputation），不作为意味着将名誉与威胁（也许是与对方本身）相脱离。因此，对双方而言，这一阶段既微妙，又复杂。尽管不作为对双方都有利，双方还是有可能无法达成共识、实现合作。

不管是威胁还是反威胁，在定义“威胁”时，我们需要格外小心。困难来自于前面已经提到的一个现实，即在威胁方已经下定决心实施威胁的时候，被威胁一方却消失了。相对于被威胁方而言，未付诸实施之前的威胁是否具有效力，很大程度上取决于威胁方在失败的情况下如何释义威胁承诺的合理性。如果威胁方的每一个枪孔都在对方的眼中历历在目，那无疑会大大削弱威胁的可信度。

为了保证威胁的可信性，威慑方必须竭尽全力隐藏自己发动威慑的真实目的和意图。但是如果一方决意要惩罚某种令人难以接受的行为，界限并非十分明显、客观的话，那么，威胁方在决定是否实施威慑的时刻，也许能够认识到自己与对方存在某个利益共同点，即双方都想避免造成彼此不愉快的行为。

为了使威慑更加准确有效，能够清楚地被对方或任何意在帮助对方的第三者所认识，威慑方有必要引入一些仲裁机制。威慑的要点在于其显著的效用性，而非显示其主体的真实意图，因此威慑与可见的行为，而非隐藏性的行为密切关联，包括主体的一些毫无意义的假动作。例如，严惩武器的携带是为了防止武器的使用；或严惩可疑行为是为了防患于未然；或惩罚非犯罪行为则是为了防止犯罪行为的发生。最后，惩罚的效果和影响必须具有显著的隐蔽

性。[1]

此外，为了在实施威慑过程中维护自己的信誉和尊严，威慑方应该注意前后事件或行为的连贯性，因为后续行为能够增强前期行为的威慑效果。如果整个威慑过程可分为若干阶段，那么这为威慑方提供了一个通过前期努力、向对方证明其威慑将继续存在的绝好机会。鉴于对方前期已经得到了一个“教训”，前期努力的效果也将愈加明显。

这一原则更多涉及威慑的程度问题。以对外援助为例，突然结束援助项目对双方来说都是一个痛苦的现实，这如同受援方不珍惜援助一样。如果接受方浪费援助的程度和提供方减少援助的速度一样（反之，接受方可能陷入孤立无援的境地并要求断交），那么，提供新援助的意愿会更加可信；否则，也仅仅留下一些教训，而不会遭受大的损失。[2]

当然，威慑方也不一定非要把整个威慑过程分为若干阶段（如果需要防范的行为和惩罚可分阶段实施）。但是这一原则至少给我们一个提示，任何超越某个界限的侵犯行为都无法令人接受。只有当被威慑的行为是具有重大影响的一系列行为的一部分时，不断增强的威胁才能比到某关键时刻完全或不完全实施的威胁发挥更有效的作用。我们甚至无法准确定义一个简单易懂、可信的“关键点”（critical point）。

为了化解对方的反威慑努力，威慑主体需要不断完善自己的威慑。有些无法分步实施的威慑行为需要剔除；尽管没有重大的实际意义，但是威慑一方可以允许一些辅助性的次要行为存在，因为它们或许将有助于威慑目标的实现。例如，威慑实施前的一些准备工作本身也许并没有重大意义，但却可以打破对方的准备计划，分散对方的注意力，从而有助于主体工作的有效开展。还有一个案例，一个踢打家犬的人应该为其第一个暴力动作负责，受到相应的惩罚，尽管平时很少有人会像他那样做。

[1] 在20世纪50年代，美国经济合作部门对外宣称，凡是认真贯彻马歇尔计划的国家都将受到援助，而那些拒绝执行该计划的国家将受到严厉的制裁。尽管援助与制裁的区别只是援助程度问题，但是由于没有奖惩的基本标准，只能依赖个人的主观评价而非客观公式，其后果是无法看到是否存在真实的奖惩以及这一计划的可行性。

[2] 贷款期间，规定贷款只能分期支付而不是一次性支付也反映了这一原则。同样的道理，对学生的一门课程要多次考察，而非一次性考察。

同样的道理，整体威慑行为阶段化只是惩罚性威胁的开始。随着时间的流逝，这一威慑行为的效果会愈加明显和严厉。如果要一个人的命不会对其产生威慑的话，那就停止向其提供食物，因为后一种方式此时比前一种方式更有效。出于对道德或公共舆论的考虑，这种方法也是留给对方一次“最后的机会” (last clear chance)。如果对方还是顽固不化，那么“最后的机会”一旦失败，被威胁方将受到公众的批评和指责。但是通常情况往往是，威慑方在其威慑行为还处于初级阶段或刚刚发挥作用时，就选择了不适宜的举动，从而无法清楚地看到最能阻止其威慑行动的拦路虎。当然，如果被威胁方一开始就知道自己正一步一步走向灭亡，那么“最后的机会”或许会真正发挥作用；如果是这种情况，威慑方或许会为对方的灭亡而非痛苦的挣扎感到自责不安。因此，这一方法能够将一次性的毁灭威慑转化为破坏性相对较小的持续威胁。同样的道理，要赶走一个房客，最好的办法不是强行将其赶走，而是给其断电断水，逼迫其主动离开。[1]

作为被威胁方，另一方也可采取化整为零的方法化解威胁风险。如果他一时无法做好全部的防御工作，也可以首先完成对其反威胁计划具有全局性意义的事项。如果威胁方的报复行动可能导致巨大损害，或被威胁方的抵抗行为也可以化解为若干阶段，那么采取化整为零的方法，也能有效化解来自被威胁方的可引起自我反应的反威胁行为。

允　诺

前面已经提到了公司的众多合法权利中的两个：诉讼权与被诉讼“权”。谁愿意被别人告上法庭呢？但是，我们必须承认被诉讼权是做出允诺的权力之一，如公司借贷、签订合同以及与一个可能的商业伙伴进行商务往来。尽管官司真正到来时，被诉讼权看起来又像是一个包袱，

[1] 这看起来与前文提到的避免引爆身上炸药和 1945 年戴高乐的部队对意大利南部省份人们所使用的策略有异曲同工之妙。 1945 年戴高乐的部队占领意大利南部省份后，警告当地人们接受纳粹援助将被看做是对盟国的敌对行为，并强行清空该地区。参阅： Harry S Truman, *Year of Decisions* (New York, 1955), pp.239~242；以及 Winston S.Churchill, *Triumph and Tragedy*,vol.Ⅵ of The Second World War (Boston, 1953), pp.566~568。

但是风平浪静时，被诉讼权却是商务往来不可或缺的前提条件。

简而言之，被诉讼权也是接受承约的权力。这里的承约是指，商业对手（也可以说是“伙伴”，但是我们更乐于将其称为“对手”）没有权力私自悔约。事实上，承约在第三方看来或是真实或是虚假。在谈判中，允诺则是向谈判对手提出的条件，这也是其中一方做出最后选择或双方谈判失控时所必需的。此外，当谈判双方达成的协议存在可能导致欺骗行为产生的漏洞时，允诺也是必要的。[1]

人们对允诺的需求并非纯属偶然，因为允诺本身具有十分重要的机制意义。人们通常很难做出一个具有说服力的、自我约束的允诺，其中包括意欲释放人质的绑架犯，甚至是人质本身也无法找到一种方式保证自己逃脱后不泄露前者的秘密。当然，如果受害人的某种行为可能受到报复或勒索，他或许会做出允诺；否则，受害人或许仅仅向绑架者保证，自己被释放后将保持沉默。尽管如此，这一可能性表明了做出允诺的困难性及其自身的重要性。如果法律不能保证价格协约的有效性；或工会无法履行自己做出的避免工人罢工的允诺；或当合约一方被起诉后没有足够的资产用于赔付时，法律不能有效保护债权人的合法权益；或没有公众关心谈判各方是否在维护各自的尊严，那么，这种谈判（至少是类似谈判）的存在也就失去了任何积极意义。

谈判不仅涉及谈判双方的“动机”结构，而且还涉及各自的收益分配，如市场垄断者也许会游说政府制定所谓的“公平交易”法或者进行相关的股票交易。如果双方试图达成协议，互不进入对方控制的市场，那么他们就必须调整各自的产品结构，避免与对方的产品结构发生冲突。同样的道理，两个国家如果都不希望对方控制某个岛屿，那么最好的办法就是想方设法破坏该岛屿的使用价值。［在现实生活中，如果“第二方承约”（second-party commitment）无法生效，人们往往采用“第三方承约”的形式。］[2]

[1] 如果威慑一方的尊严或荣誉与对方紧密联系，威慑行为也可能成为一种允诺；但是，一旦对方单方面化解威慑行为，这种威慑就不是允诺行为了，因为他无法肯定地将对方对自己的未来预期与对方的行为脱离关系。

[2] 在早期，人们通常互相交换人质以示诚意。

尽管某些承约的履行有时不容易被公众察觉，但是如果承约发生在下面场景中，如人们在无记名投票中出卖自己的投票权、政府允诺企业向议会建议某项立法、雇员保证不从仓库中私拿产品、老师保证在教室外保留自己的政治观点，以及一个国家许诺努力促进出口，人们或许能够采取某种有效的监督方式或衡量尺度。明确的谈判结果往往受到许多因素的制约，但是其中只有一个因素能够真正反映谈判双方所达成的协议。因此，谈判必须以人们能够监督的方式进行，尽管这并非谈判所刻意企求的。例如，一位候选人在选举中脱颖而出，那么，他就必须回报接受其贿赂的投票人，不管后者采用何种方式；或者给企业家的佣金通常以业绩为标准，而非其工艺和努力程度；或者奖赏警务人员的标准是犯罪率而非其敬业精神；或者因一个职员的过错而惩罚所有职员。当衡量业绩的标准涉及程度问题时，谈判双方必须制定相关的衡量标准以定义作为和不作为，如仓库中丢失某件物品就是偷窃行为存在的证据；某项出口的增长就是“充分”(adequate)努力的成果。由此可见，局部成绩完全可以说明整体成绩的存在。[1]

化整归零的阶段化方法不仅适用于威慑，同样也适用于允诺。只有当谈判当事人认识到，双方由于缺乏互信或因为一方的收益大于谈判得到的暂时利益而忽视了达成共识的机会时，多种协议方式才能发挥效用。双方必须坚信，谈判对手不会以牺牲未来的机遇来破坏双方在谈判开始阶段营造的互信气氛。然而在通常情况下，谈判选手很难建立这种互信。此外，将谈判过程阶段化也是为了创造良好的互信气氛，因为在整个谈判过程中，谈判双方在任何关键问题上都不会轻易相信对方的审慎行为。但是，如果谈判过程可以被划分为若干阶段，那么双方就会愿意以较小的风险营造互信的气氛，目的是为了向对方展示建立互信的诚意，并希望对方也能展示其应有的合作诚意。因此，在就关键问题进行谈判之前，谈判双方可以先就一些次要问题进行“练习”(practice)，以便建立互信的诚意和培养长期合作的精神。

[1] 无法做出有效的允诺或履行相关的责任、都有可能保护当事人免遭外来威胁。强制性的无记名投票对那些乐于出卖自己投票权的人也许不起作用，但可以保护当事人免遭恐吓。

尽管历史不会重演，我们仍然可以通过将谈判过程阶段化，实现谈判的连贯“等效性”(equivalence)。例如，双方都以对方为前提条件，同意各向红十字组织捐献 100 万美元。假设一方先捐献了，那么另一方就有可能不履行自己的义务；与此同时，双方对彼此的互不信任也将阻碍双方履行各自的允诺。但是，如果双方同意分阶段履行自己的捐款承诺，那么一方就会通过观察对方的每次履约情况来评价其信誉。而且，鉴于一方会时刻监督另一方的第一次履约情况，则没有人愿意以一时的欺诈行为而丧失长久的信誉。最终，这一动机结构的调整大大降低了谈判初期阶段的风险性，也让谈判主体清楚认识到了诚信的重要价值和作用。

除此之外，谈判选手就次要问题进行先期谈判还有另外一个好处。一般情况下，只有当至少一方谈判主体提出谈判动议时，双方才有可能进行谈判。但是，如果一方表现出过高的谈判意愿也许会阻碍谈判动议的提出；而且，谈判双方总是期望另一方主动迎合自己的立场，因为先前成功的谈判经验告诉他要避免表现出过分的谈判意愿。

博弈游戏

虽然谈判涉及多种因素，但无一例外地涉及到谈判条件、威胁、允诺和信息沟通的问题。在下面这个博弈游戏中存在多个变量，可以有效地帮助我们理解前面的问题。游戏中，游戏选手双方都有两个可供选择的变量：上北方向的 A 或 B 和右东方向的 α 或 β。一个方向上的变量与另一方向上的变量组合不同，产生的结果也将不同，即：AB、$A\beta$、$\alpha\beta$ 和 αB 这意味着双方在北和东两个方向上都会有所得，亦有所失；而且一个方向上的所得不能弥补另一方向上的所失。一般而言，一个选手的选择很大程度上取决于另一选手的选择。

该游戏被量化为垂直方向为北、水平方向为东的二维图表，四个选择组合分别表示为 AB，$A\beta$，$\alpha\beta$ 和 αB。这个游戏看似简单，实际上游戏中存在多个变量，并受到多个因素的影响，如四个选择组合的相对位置、选手的游戏规则、沟通的可能性、条件的提出方式、允诺的有效性，以及如果选手同时参与两个以上游戏，游戏是否可以合二或合多为

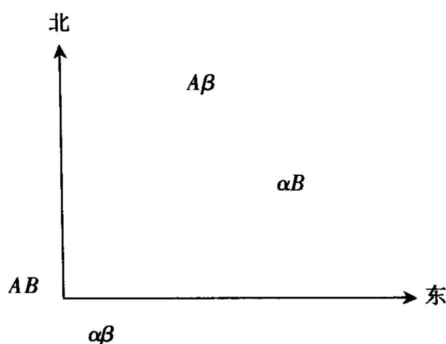


图 2-1

一。根据选手对对方赋予每个变量的“值”(values)的预期判断，异向变量可以自由组合。为了易于操作，这里我们以一个简单方式将 8 个“值”分别向大家展示。此外，我们还排除了补偿因素和博弈游戏之外的各种威胁行为因素。下面是一些博弈游戏的案例分析。

图 2-1 表示的是一个大家比较熟悉的“普通”谈判情景。在此情景中，我们可以假设在开始游戏之前，选手双方已经达成了某种共识。$A\beta$ 和 αB 可被看做是二者能够达成共识的选择性结果。起点处的 AB 和 $\alpha\beta$ 被双方赋予零“值”，并被看做是谈判中的“无货”（No Sale）状态。我们规定，第一个做出选择、提出条件的选手则为获胜方。如果北首先选择了 A，他将得到 $A\beta$，留给东的选择项是 $A\beta$ 和 AB。但是，这里我们需要提醒大家注意的是，$A\beta$ 显然是东的最佳选择。如果东首先选择了 B，则北将被迫从 αB 或退出游戏（即 αB 或 AB）中做出选择，最终也只好选择 αB。实际上，这里的第一选择权相当于游戏中进行“第一步”(the first move) 的权利。在选手数量相同、交替选择的类似游戏中，选择“第一步”的选手往往掌握游戏的主动权。如果双方同时都做出了错误的选择，即北选择 A，东选择 B，那么游戏将会出现僵持局面。

图 2-2 表示在威胁存在情况下的博弈情景。我们将 AB 看做是现状，北现在计划向 α 移动（目标为 αB）；与此同时，如果东愿意的话，他也可选择向 β 移动（目标 $\alpha\beta$)，对北形成威胁。如果北率先采取行

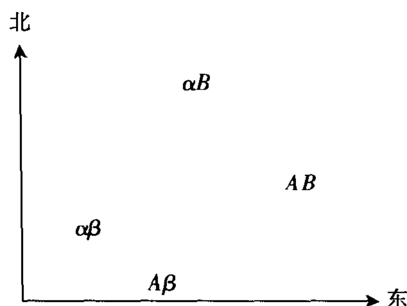


图 2-2

动，东将被迫向 β移动；同样的道理，北可以在东采取威胁行动之前向 α移动。但是，如果东能够有效地威胁到二者都不感兴趣的 $\alpha\beta$，那么北只好被迫从 $\alpha\beta$ 和 $A\beta$ 中选择 $A\beta$，除此之外，北别无选择。请大家注意，这里东没有必要像在图 2-1 中那样率先做出选择；东在此只需要根据北的选择 A 或 α，做出自己相应的选择 B 或 β。如果东首先选择，他惟一的优势在于获得选择"第一步"的权利。但是在这里，只要双方交

替选择，无论谁先走"第一步"，北都将赢得 αB。(若东选择 B 而非 β，将迫使北选择 AB 或 αB 中的 αB，而非 $\alpha\beta$ 或 $A\beta$。若北首先选择，北将选择 α 而非 A，从而迫使东选择 $\alpha\beta$ 或 αB 中的 αB，而非 AB 或 $A\beta$。)

图 2-3 表现的是一个允诺场景。无论双方谁先选择第一步或是否同时选择第一步，βB 都将作为"极小极大"点存在。虽然双方都有机会到达该点，但谁也不能以相对不利的条件威胁另一方。 $A\beta$ 与 αB 相

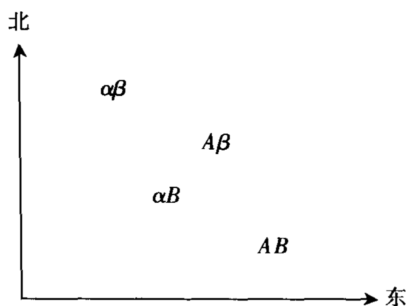


图 2-3

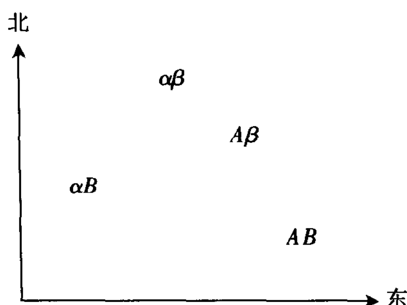


图 2-4

比，双方都希望得到前者；但是要实现这个目标，双方之间需要建立互信或相互达成有效的允诺。但是无论谁选择第一步，另一方都会有实施欺骗行为的动机。如果北选择了 A ，东将得到 AB；反之，如果东选择了 β，北将得到 $\alpha\beta$。如果双方同时做出选择，任何一方不仅自己希望采取欺骗手段，而且还会假设对方也会采取某种欺骗手段。无论是出于故意，还是为了保护自己免于对方的欺骗，只要有一方采取了欺骗手段，双方选择的结果都将是 α或 B。而且，至少有一方必须弃权，而另一方则获得“第一步”的权利。如果双方同时做出选择，那么二者必须达成某种有效的允诺。

图 2-4 与图 2-3 基本相同，除了前者中 αB 的位置向左偏移。此外，在缺乏沟通的情况下，无论双方谁先选择或是同时选择，北都将赢得 $\alpha\beta$。但是，如果东能够与北进行有条件的沟通，他将迫使北选择 A 并得到 $A\beta$。但是这一条件很大程度上不仅仅是允诺或威胁的问题，或许允诺和威胁二者兼而有之。如果北选择 α，东将以 αB 威胁北；或者假如北选择 A ，东将允诺不占领 AB。但是，如果东仅仅对北进行强制威胁，还无法达到强迫北放弃选择 α 的目的，因为相对于 AB 而言，αB 对北更具有吸引力。如果东能够自由选择 B，那么北将选择 A ，并得到 AB。为了得到 A 或 α ，东必须做出选择；否则，只能从 AB 中退出或在 αB 处被消灭。

如果对图 2-5 和 2-6 所表示的博弈游戏分别进行分析，我们将得不到任何有价值的结论。但是，如果将二者放在一起进行综合考虑，我们

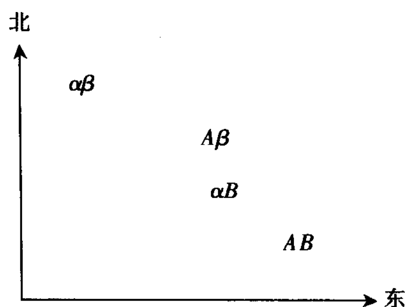


图 2–5

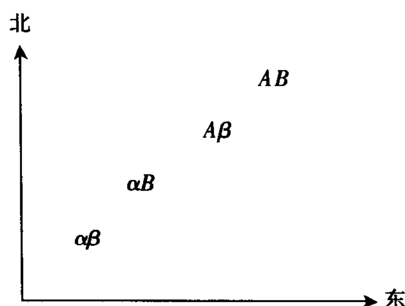


图 2–6

会得到一个敲竹杠式的威胁场景。在图 2–5 中，αB 处存在一个极小极大的解决方案，双方都可能得到 αB；但是，他们都无法取得更好的收益，并且不需要任何形式的合作和威胁。在图 2–6 中，尽管双方间的利益分布不同，但同样不需要任何合作和威胁。无论沟通或行动指令存在与否，双方的收益都是 AB。

现在，我们假设两个博弈场景被同时列入综合考虑的范围，游戏选手依然是原来的两个选手，假设双方都可以通过威胁对方来改善自己的处境。例如在图 2–6 中，东可以威胁选择 β 而非 B，除非北在图 2–5 中选择 A 而非 α；反之，北在图 2–6 中可以威胁选择 α，除非东在图 2–5 中选择 β。假设在图 2–6 中，双方之间存在充足的思考时间，威胁也发挥了极大效果，那么威胁方在图 2–5 中将取得最终胜利，在图 2–6 中也不需要付出任何成本，因为其威胁承诺已经达到了预期效果，无须再付诸实施。因此，威胁方在图 2–5 中取得胜利后，又在 2–6 中如愿以偿地

得到 AB。换句话说，图 2–6 表示的博弈场景中出现了此前被我们排除在外的某些因素——某种存在于游戏以外的威胁行为。以图 2–5 所表达的语境来看，图 2–6 完全是多此一举，因为东完全可以威胁北，如果后者不选择图 2–5 中的 A，就将北的房子化为灰烬。但是事实上，类似的敲竹杠式威胁很少发生，因为这些威胁行为大多需要一定的场合、目标及方式，而且经常被认为非法、不道德和顽固不化等。当一个完全无理的威胁行为无法奏效时，选手若选择以同一个议程对两个问题进行通盘考虑，也许会成功。

如果北无法采取必要的威慑行为，就只得努力避免受到东的威胁，这时沟通不畅对其非常有利。即使沟通畅通，只要两个谈判尚未以同一个议程进行，形势对北还是十分有利的；或者即使北无法避免与东就一个问题进行沟通，只要他将不同问题分别按照不同的议程进行下去，并且要求是否给予补偿则要根据谈判的结果对其是否有利来决定，那么，形势对北也是比较安全的。如果北能够迫使对方先进行图 2–6 中的游戏，并避免对威胁做出反应，他同样可以保护自己免于来自东的威胁。此外，在图 2–5 中，只要北能够在威胁来临之前做出自己的选择，就也可以避免受到威胁的伤害。但是，如果北在图 2–5 中做出了自己的选择，而且图 2–6 先于图 2–5 开始进行，并且东仍然能够威胁北，那么，除非北首先选择图 2–5 中的 A，否则仍将被迫选择图 2–6 中的 β。在这种情况下，北的优先选择权将转化为不利条件，因为该选择权迫使北在图 2–6 开始之前，首先进入图 2–5 中。

巧合的是，如果将图 2–2 中的 AB 位置在垂直方向置于 $\alpha\beta$ 之下，那么我们将得到一个非常重要的原则，即：以对北“不利的”方式移动某个点，实际上却可以产生对北有利的结果。在图 2–2 中，阻碍北取胜的威胁因素很大程度上来自于 AB 相对于 $\alpha\beta$ 对北的吸引力；如果 AB 相对于 $\alpha\beta$ 对北不利，那么他就可能化解威胁；如果没有威胁存在，北将一定赢得 αB。有关这一原则的抽象案例有力地证明，谈判中的劣势完全可以转化为优势。

第3章

谈判，沟通与有限战争

有限战争要求具备某些有限条件。如果战争要具有稳定性，那么战略行为也将如此。但是，有限条件要求战争双方之间至少存在某种程度的共识或默契。然而，这种共识或默契通常很难达到，原因不仅仅在于现实中存在很多不确定因素和双方利益的尖锐冲突，更重要的是在战争期间，甚至在战争爆发之前，双方已很难进行有效的谈判活动；或者由于战争的存在，战争双方无法进行有效的沟通。而且，有时战争一方为了增加战争对另一方的威慑力，往往不愿意主动就有限条件进行相关谈判；甚至一方或双方都担心自己的主动谈判意愿有可能被对方释义为妥协。

因此，对默式谈判 (tacit bargaining) ——谈判双方信息沟通不完全和无效情况下的谈判模式——的研究具有非常重要的意义。默式谈判涉及的问题往往非常广泛，诸如有限战争或与此有关的问题，或有限竞争问题、司法管辖权的问题、交通堵车时的博弈情景，甚至包括如何与一个从未打过招呼的邻居相处等等。问题在于当一方或双方无法或不愿进行显式谈判（explicit bargaining），或双方都无法相信会与对方达成任何明确协议时，谈判双方如何实现暂时的妥协。在第 3 章中，我们将检验默式谈判所涉及的一些基本概念和原则，并努力得出一些有关有限战争或类似情况的说明性结论。同时，本章将努力得出这样一个结论：默式

谈判具有的原则同样有助于我们理解信息沟通完全畅通和有效的显式谈判，尽管二者存在逻辑上的不同。

最有趣也是最重要的场景是，双方之间存在利益上的冲突。为了满足研究的需要，我们有必要从简单的场景开始，即在无法进行有效沟通的情况下，双方或多方之间存在共同利益，并且面临的问题不是如何调解利益关系，而是如何为实现共同利益来协调彼此行为。这一特例将向我们揭示解决利益冲突的默式“讨价还价”所必需的一些原则。

默契（利益共同）

假设一对夫妇在拥挤的百货商场失散，事先也没有约定见面的地点，但是他们还是有可能找到对方。也许一方一直认为，对方也希望在一个双方都认为比较醒目的地点与自己会合，因为夫妻双方都认为该地点比较醒目，易于发现对方或被对方发现。而且，一方不会轻易判断对方首先要去的地方，因为在上述情况下，对方首选的地方可能也是其所希望的。换言之，无论发生什么情况，一方所到之处都将是对方所期望的地方。我们可以如此不断推理下去，一方所想的问题不是“如果我是他，我该去什么地方呢？”而是“如果我也像他一样在思考同样的问题‘如果我是他，我该去什么地方呢？’我该怎么做呢？”类似的情景需要双方间的相互默契，对同一场景提供的信息进行同样解读，并努力促使双方对彼此的行为进行相同的预期判断。双方必须认同（mutually recognize）某种能够协调彼此对对方行为预期判断的暗示符号。当然，我们既无法肯定他们一定会重逢，也不能肯定双方一定会对同一暗示符号进行相同的解读。但是，夫妇双方如以这种方式寻找对方，成功的机率一定比他们盲目地在商场里瞎转要高得多。

读者可以尝试着解决下面图示（图 3–1）所出现的问题。两个空降兵意外地降落到了图中所示的地区。双方都随身携带了一张地图，并知道对方随身也有一张地图，但不清楚对方降落的具体地点，也无法与对方进行联络，双方只有尽快地找到对方才能获救。那么，他们是否会研究随身携带的地图，并协调彼此的行动呢？他们是否会不约而同地想到并在地图中标示的某个地点会面呢？

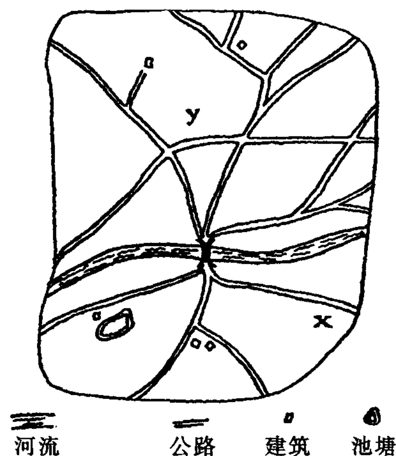


图 3-1

作者认真研究了这个问题以及一份缺乏科学性的调查问卷中出现的一些类似问题，得到的答案是肯定的：人们通常会协调彼此的行为。下面所列的抽象难题也属于大多数人试图解决的问题范畴。当然，问题答案的评判标准是，只要每个人认为自己的答案具有合理性，那就是“正确”的答案。读者可以先回答下面的问题，然后参照脚注测试一下自己的能力[1]。

1. 请在“heads”和“tails”二者中选一个，如果你和同伴的选择相同，你们都将获胜。

[1] 作者进行的一份调查问卷结果显示，问题 1 中，36人选择了“heads”，只有 6人选择了“tails”。问题 2 中，41 人中的 37 人选择了前 3 项，7略微领先于 100，13 位于第 3位。问题 3 中，41 人中的 24 人选择了左上角的空格。剩下的 17人中，有 14人分布在同一对角线上。问题 4 中，大多数人选择到中央车站（服务台）会合，实际上，他们都在中午 12 点实现会面。问题 6 中出现了不同的答案， 40%的人选择数字 1。问题 7 中，41 人中的 12人选择 1 000 000美元，只有 3人选择 10 美元，其中 2人是 64 美元，不要小看这个数字，这相当于今天的 64 000美元。问题 8 没有什么问题，41人中的 36人的选择各占一半。问题 9 中，22人中的 20 选择了 RoBinson。在另一个公式中，Jones 和 RoBinson 以 28 票位于第一。作者这么做是为了验证不分胜负的情况下，要取得一致如此困难。但是，被调查者都克服了这个问题， 18人中的 16人选择了 Jones（主要原因是 Jones 所在的位置有利），虽然解决了主要问题，但也忽视了相关的次要问题。在图 3-1 中，8 人中的 7人选择在桥上重逢。

2.请选择下列数字中的一个，如果你们的选择相同，你们都将获胜。

7　100　13　261　99　555

3.请选择下面的方框　如果你和同伴的选择相同，　你们都将获胜。

□	□	□
□	□	□
□	□	□

4.你将在纽约与某个人会面，但没有人告诉你会面的地点。你与对方也没有事先约定会面的地点，而且无法与对方进行联络。你和对方只能自己寻找会面的地点。你将怎么做？

5.在问题 4 中，你只知道会面的日期，但不知道会面的具体时间。你和对方必须设法约定时间在问题 4 中的地点会面。你将怎么做？

6.写下你认为吉祥的数字，如果你和同伴写的数字相同，你们都将获胜。

7.请写下一笔金额。如果你和同伴写的金额相同，你们都将获胜。

8.你被要求将 100 美元分成两部分，分别标为 A 和 B。你的同伴也会像你那样做。如果你们的结果相同，你们每人都将得到 100 美元；反之，你们将一无所获。

9.在第一个投票中，候选人的票数如下：

Smith……………………19　　Robinson……………………29

Jones……………………28　　White…………………………9

Brown……………………15

随后将进行第二次投票。除非你期望自己的候选人获得了多数票，并为此得到回报，否则你绝不会关心投票的结果。同样，所有的投票人都希望自己所选的候选人获得多数票，而且大家都知道这是每个投票人的心愿。第二次投票时，你将把票投给谁呢？

虽然上面的问题都是人为设计的，但是这些问题向我们揭示了一个道理：人们通常只有在得知别人将做出和自己同样的行为时，才会与他

人产生共鸣，达成某种共识。大多数情况下——对那些参与此类游戏的人来说，也许每一种情况都是——都会出现某些合作的契机，如某个“关键点”（focal point）使双方成功地对彼此预期做出判断，从而达到某种默契。能否找到问题的最佳解决方法或是任意一种解决方法——只要是双方认同这种方法是他们所要寻求的答案（*key*）——很大程度上取决于双方的直觉而非逻辑思维推理，或许依靠来自双方对相似事物之间的类比经验、先例、偶然巧遇、对称性、审美观或几何原理、诡辩推理，以及当事人的自身条件和对彼此情况的了解。奇思妙想也许会帮助一对夫妇“分而重逢”；理性的逻辑推理也可能会使一方回忆或期望另一方回想起来他们以前处理类似紧急情况的方式方法。我们不敢保证他们一定能找到问题的答案，但是他们这样做的可能性远远大于抽象的随机率。

解决这一问题的众多“解决方法”具有一个非常显著的共同特点，即这些线索或关键点都具有某种显著特征，易于发现。但是，这些显著特征随着时间、地点以及当事人的变化而变化。大多数普通人在一个环绕的圆形区域走散后，通常都会“自然而然”（naturally）想到在圆形地带的中心区域与对方会合。但是，如果是在一个非常规形区域失散，那就只能依靠个人的方位感在该区域的重心地带与对方会面。作者曾经以多个地图进行实验，结果证明：如果一个地图标有多个住宅和一个十字路口，人们大多会本能地趋于十字路口；反之，如果一个地图标有一个住宅和多个十字路口，人们大多会本能地趋于住宅。这充分说明，惟一性能够产生独特性，从而吸引人们的注意力。但通常情况是，惟一性在避免似是而非上具有更重要意义。在地图上，或许住宅相对于其他标识，更容易引起人们的注意，但是若地图上同时存在三个住宅，并且具有相似的标识，那么人们在某个住宅会合的可能性或许只有三分之一。认识这一现实有助于我们避免在解决问题时，将住宅看做一个显著的“线索”。[1]

[1] 正如作者的地图试验所证明的，这是“正确”的逻辑推理过程。在标有多个十字路口和一个住宅的地图中，选择住宅的 11 人胜利实现会合，但是选择不同十字路口的 4 个人没有一个人实现目标。

在最后的分析部分，我们讨论了想像力问题和逻辑推理问题。就其本质而言，逻辑推理具有很强的诡辩性。在上面的游戏中，诗人可能会比逻辑学家做得更好，因为与其说这样的游戏是国际象棋，还不如说是充满字谜的文字游戏。逻辑有时也会发挥积极的作用——例如在问题 6 中，大多数人趋向于数字 1 都是基于逻辑推理。但是在大多数情况下，主体在进行逻辑推理之前，必须依靠想像力从当前环境提供的众多具体细节中发现主要线索。

默式谈判（利益冲突）

假设空降兵双方都不喜欢走路，那么我们在解答问题的时候，就必须考虑双方利益的冲突因素。如果双方能够进行有效的沟通（当然这里是不允许的），那么双方可以就会面的地点进行商讨；而且每一方都会根据自己的喜好，选择一个靠近己方的地点或是会合场所。而在缺乏有效沟通的情况下，双方则会抛开利益的分歧，努力在彼此间寻求共识。如果某个“醒目”的地点引起他们共同的关注，而这个地点更接近于其中一方的期望，那么他就是这场博弈的胜者。即使这一结果对另外一方多么不利，后者也无法提出任何异议，而只能接受这一现实。地图本身提供的谈判“动案”（proposal）——假如事实上，确实发挥了作用——显然是惟一可行的发盘（offer），因为在缺乏有效沟通的环境中，双方之间根本不存在讨价还价的空间。在协调沟通压倒一切的情况下，作为一个副产品，一场冲突就可能被避免——我们也可以认为被双方忽视了。

鉴于双方都失去了在有效沟通条件下应该得到的东西，用“赢”和“输”来形容游戏双方似乎并不太准确。如果双方降落的实际地点相距很近并远离住宅，那么他们就没有必要长途跋涉到住宅会合。他们只需要辨明自己所处的位置，到位于二者中间的某个地点会合；或者一方“获胜”，另一方的所失大于前者的所得，如双方降落于住宅的同一侧，虽然二者经过长距离跋涉——距离远远大于理想路径——终于实现了会合，但是与双方商讨后得到的结果相比，距离住宅较近的一方仍然赢得优势。

最后一个案例表明，对谈判某一方而言，缺乏有效沟通也许并不是什么坏事。如果一方能够准确判断自己的优势，并对自己能够预见的谈判结果充满信心，那么他就可以设法破坏沟通渠道或拒绝提前为相关问题与对方进行任何有效的协商。作者设计的问卷中有一种情况， A 知道 B 的位置，但是 B 不知道 A 的位置（除此之外，双方掌握等量的信息）。处于 B 情景的大多数答卷人表露出毫不在乎、镇静自如，并乐于接受无知的状态。而处于 A 情景的答卷人则几乎异口同声地表示自己将别无选择地前往 B 所在地。这表明在一定条件下，能够发出信息却无法接受信息，对一方而言将是一个优势。例如，一方将自己所处的位置告诉另一方，并声称自己的发报机只能发送信息，不能接通信息，即自己只能原地待命，等待对方到来。结果可想而知，另一方别无选择，只能接受对方的条件，因为他无法与对方进行有效的沟通，对方收不到任何信息。[1]

作者进行了大量有关利益冲突博弈问题的问卷调查。尽管有些问卷甚至涉及不公平的博弈场景，但总体而言，问卷得出的结论基本上与完全合作博弈场景中的结论相符。这些博弈场景均要求答卷者进行有效的协作，并提供了可能导致双方利益不同组合的选择项。在提供的众多选择项中，一些选项似乎是双方达成共识的必要条件，而且处境相对不利的一方往往也并不在意，因为他知道对方期望自己采取这样的行为。由此可见，在缺乏有效沟通的情况下，双方之间一定能够达成某种共识。这些博弈情景的共同之处在于，任何一方都无法通过雕虫小技赢得对方。只有当一方完全按照对方的条件做出选择时，他才可能被认为是输者。双方既是二者间达成共识的受害人，也是其受益人。无论如何，一方都不能拒绝对方提出的条件。因为达成共识的迫切性远远大于利益的分歧，双方只有同对方合作，才能取得双赢，否则只能两败俱伤。其中有些问题只是在前面问题的基础上略作修改，正如我们针对地图问题所假设的：长途跋涉是徒劳无功的。

1. A 和 B 在无法沟通的条件下，必须在“ heads ”和“ tails ”之间做

[1] 第 2 章对类似看似矛盾的案例进行了详细的论述，即在谈判中，一般意义上的劣势也可能转化为优势。

出选择。如果二者都选择“heads”，A 将得到 3 美元，B 则得到 2 美元。反之，如果二者都选择“tails”，A 将得到 2 美元，B 则得到 3 美元。如果两人的选择不同，二者将一无所获。假设你是 A（或 B），你将如何选择？（注意：双方的选择具有随机性，输赢机率参半，并且只能得到 1.25 美元。）

2. 你和另外两名同伴（或对手）分别从 A、B 和 C 中任选一个。每个人以任何顺序写下这三个字母。如果三个人的排序顺序相同，三个人将共获 6 美元，具体分配方法为：第一个完成排序的人将得到 3 美元，第二个人将得到 2 美元，剩下的 1 美元归最后一个人。如果三个人的排序不同，都将一无所获。假设你是 A（或 B，或 C），请写下你的排序：

3. 你和你的同伴（对手）中的一人将得到一张空白纸，另外一人将得到一张写有字母 X 的纸。每个人都有两个选择：持写有字母 X 的纸的一方可以不必理会 X 或将其抹掉；另外一人保持纸张空白或在纸上写上字母 X。你们在没有沟通的情况下做出选择，如果你们之间仅存在一个 X，那么持有 X 的一方将得到 3 美元，另外一方将得到 2 美元；如果存在两个 X 或两张白纸，双方将一无所获。现在，你得到写有字母 X 的纸，你将如何处理？（反之呢？）

4. 你和你的同伴（对手）现在拥有 100 美元，要求双方在缺乏沟通的情况下决定分配的方法，每人将自己希望得到的金额写在一张纸上。如果两人的期望金额加起来不超过 100 元，每人将得到自己所写的金额；反之，双方将一无所获。你希望得到______。

5. 你和你的同伴从下面的字母中任选其一。如果选择了相同的字母，你们都将获得奖金；反之，双方将一无所获。所获奖金金额大小取决于你们选择的字母，双方所得的奖金金额不等，而且金额最大的字母不一定给你带来最高奖金。你们的奖金如下：

K…………………………4 美元　L…………………………2 美元
G…………………………3 美元　R…………………………5 美元
W……………………1 美元

你不知道对方的想法。你首先向对方建议对自己最有利的字母 R。

在对方做出选择之前，裁判将警告双方以下时间不得进行沟通，否则双方出局。你必须迅速选择自己的字母，同时希望对方选择同样的字母。那么，你会选择哪个字母？（第 2 种情况是，5 个字母代表的金额分别为：K–3 美元，G–1 美元，W–4 美元,L–5 美元，R–2 美元。在裁判发出警告之前，你的同伴首先向你建议选择字母 R。)

6. 两支敌对军队分别位于 X 点和 Y 点，标示二者方位的地图类似于图 3–1。现在，两军的指挥官要求自己的部队尽可能多地抢占土地，并知道对方将采取同样的行为。但是，双方都竭力避免接触交火，而且深知对方也具有同样的意图。双方命令自己的部队在预定范围内行动，如果遇到抵抗将进行反击。但是一旦双方部队开始行动，形势很大程度上将依赖于双方的指挥官划分的界线。倘若二者划定的界线重合，双方部队必将发生冲突。若在双方占领的重叠区域存在尚未被占领的有利地形，这时形势将更加不稳定，一场冲突在所难免，除非二者占领的区域不存在重合现象，且在双方之间存在一定的空地。只有这样，双方才能真正将抢占的领域据为己有。就占领的土地和设施而言，谁具有相对价值优势，谁将赢得主动权。现在，你所统领的军队的位置为 X（或 Y)，请在地图上标出你希望占领的土地范围。

7. A 和 B 的年收入分别为 100 美元和 150 美元。双方都知道对方的收入情况，并被告知以后将交纳个人所得税共 25 美元。如果双方能够达成分配二者共同年收入的协议，他们将以达成的方案分摊双方的个人所得税，前提是双方不得进行任何沟通。现在双方将自己希望分摊的税费比例写在纸上，如果双方希望分摊的税费加在一起等于或大于 25 美元，双方将以自己希望的比例交纳税费；反之，如果双方希望分摊的税费加在一起小于 25 美元，每人将交纳 25 美元。如果你是 A（或 B)，你希望分摊

8. A 和 B 同时合租一个房子。B 捡到了 A 丢掉的一部分钱。根据房东契约，除非 A 向对方给予一定报酬，否则 A 不能索要 B 捡到的钱；反之，除非 A 同意，否则 B 也不能将捡到的钱据为己有。如果双方无法达成协议，B 捡到的钱将归房东所有。 B 捡到的钱数为 16 美元，A 提出给 B2 美元作为条件。但是 B 拒绝 A 提出的条件，并提出分摊所捡

钱数的一半，即 8 美元。双方为此争吵不休，并引起房东的干涉。房东提出建议，A 和 B 一次性地将其所希望得到的金额写在纸上，并不得进行沟通。如果双方希望的金额加起来等于 16 美元，每人将得到自己所希望的金额；反之，如果双方的金额大于 16 美元，该笔钱将归房东所有。现在，双方坐下来冷静考虑自己希望得到的金额。这时，一位知名的、受人尊敬的调解人出现，并提出自己的建议。他表示，自己不会涉入谈判之中，只会提出一个“公平”的建议。他走向 A，并对其说：“在现有条件下，最好的分配方法是二一开，即失主得到全部金额的三分之二，拾到者得到剩下的三分之一。我同样也会向 B 提出同样的建议。”后来，他果真履行了自己的诺言，将其对 A 所说的话对 B 重复了一遍。如果你是 A（或 B），你希望自己得到多少？

作者的调查结果如脚注所示。[1] 在那些包含“你”和“他”或“A”和“B”之间存在不对称性的问题中，A 的模式和 B 的模式在分配收入的问题上相辅相成。正如脚注 1 所表明的，最终的结论是：在上面的大部分情景中，当事人都能够达成某种共识，妥善解决彼此间的问题。在允许任意选择存在的情况下，双方理性的选择结果总是优于盲目选择的结果。甚至是最不公平问题中的弱势一方，也不得不接受问题中有关合作协议的约束。

[1] 问题 1 中，A 组的 22 人中有 16人和 B 组的 22 人中的 15人选择了 heads。如果知道 A 的结果，heads 将是 B 的最佳选择；同样，如果知道 B 的结果，heads 将是 A 的最佳选择。综合考虑，双方理性选择显然优于盲目选择。当然，如果双方都希望得到 3 美元，那么二者都将一无所获。问题 2 除了结构上比较紧凑，逻辑上与问题 1 基本相同。A 组的 12 人中有 9 人，B 组的 12 人中的 10 人以及 C组的 16 人中的 14人成功地选择 ABC（剩下 7 人中的 5 人失败）。问题 3 的结构基本上与问题 1 相同。A组的 22 人中的 18 人与 B 组中 19 人中的 14 人达成共识，A 得到 3美元。问题 4 中，40 人中有 36 人选择 50 美元。（剩下 2 人分别选择 49 美元和 49.99 美元）。问题 5 中，首先提出条件的 8人中的 5人选择和另一方的 9人中的 8 人选择 R。问题 6 中，X 组的 22 人中的 14 人和 Y 组的 23 人中的 14 人选择沿着河边划定界线。应该引起注意的是，另外跨过河界的 15人出现了 14 条不同的界线。结果，出现 8×7 种结果，其中只有 1种结果成功。这强调了这个解决方案的“正确性”。问题 7 中，选择年收入为 150 的 6人中有 5人和另一组的 10 人中的 7 人达成税费分配方案为 15:10 。问题 8 中，丢钱的 8人和拾到钱的 7 人同意调解人的建议，接受 5美元的报酬。

当然，上面每个问题的答案"线索"也各有不同。尽管传统俗约没有如此强的约束力，但是根据传统习惯，头显然应该排在尾巴的前面，类似的情况还有 A、B、C 字母排序。很显然，最初的 X 一方之所以能够战胜白纸一方，主要因为双方认为"现状"显然不会发生很大变化；字母 R 能够取胜的原因在于其与第一方提出的条件相吻合。原则上讲，道路和小河能够发挥同样的作用，它们之间的差异也能允许随机条件的存在。更准确地说，标有多条道路的地图并没有指出哪个路径最好，因此只能放弃小路而选择具有惟一性的显著性标志物——小河。（也许在同一地形相对对称的地图中，得到的结果可能与分钱问题的答案相同——5-5 方案。但是，地图的非常规性排除了几何方法。）

税费问题揭示了在分钱问题上咨询建议的强大威力，这个问题与如何分配 100 美元的问题在逻辑结构上基本相似。事实上，这个问题应该这么表述：每方每年交纳 25 美元税，而且根据双方达成的分配协议，二者可以分配返还给双方的 25 美元。这样一来，该问题在逻辑上与问题 7 相同，与问题 4 也仅仅在于数字上的不同：这里是 25 美元，而在问题 4 中是 100 美元。通过突出收入金额在问题中的相关性和重要作用，收入金额的变化改变了分配方案的关键点，使最终的分配方案为 10-15 而非 12.5-12.5。如果收入在这个问题中发挥了重大的作用，那么，当分度税率的条件存在时，为什么完全比例税如此明显？答案一定是，在缺乏沟通的情况下，毫无疑问，没有哪个分度税率分配方案如此清晰易懂。如果双方无法进行沟通，在默认条件下，双方只能采取简单易懂的比例分配方案了。首先，收入的存在排除了对半分配的可能性。其次，在双方无法有效沟通的情况下，简单易懂的 10-5 比例分配方案是惟一具有可行性的方案。问题 7 中，通过安排诸如家庭规模、消费习惯等辅助因素，向我们验证了一个相同的原则。这里，收入比例分配的作用显然已经被削弱了。高收入者和低收入者此前就纳税问题达成的优势方案现在变成了简单的 5-5 分成。修改后的收入比例分配标识遭到了大家的一片唏嘘，被选择的竟然是粗制滥造的平等分配标识。

最后，问题 8 与问题 4 具有相同的逻辑结构。在问题 8 中，如果双方所希望的金额总数不超过 16 美元，他们就能得到各自希望的金额。

但是，在机制安排上存在歧视性和不公正。无论从法律的角度还是从道德的角度，捡到钱的人和丢钱的人之间的关系都极不平等，所以双方不可能达成“5–5”分配方案。调节人的建议只能说明存在其他分配方式的可能性，其作用也只是表现在圆滑地提出 11–5 的分配方案，这也是大家能够普遍接受的方案。

在上面的案例问题中，仲裁因素在某种程度上，无不对各个问题的结果产生了不同程度的影响。无论是从旁观者的角度还是从当事者的角度，这些结果也许并非完美无瑕、公正无疑。即使是“5–5”分成的分配方式，也许只是纯粹数学意义上的公平。即便如此，这种分配方式也并非如我们想像中的那样绝对公平，因为我们无法以绝对精确的数据和尺度衡量其公正性，如资金的来源、对方申诉人提出的相关要求或是其他法律或道德上的索赔依据。尽管争论勒索和绑架之间的区别本身就不公平，但是这在问题 4 中却具有数字上的意义。

如果我们要问是什么因素决定了上述问题的答案，回答还是协作因素。每一个问题都需要当事人为了实现共同利益，寻求彼此间的共识，尽管在达成共识的过程中会遇到很多阻力。但是，在众多可能的选择机会中，通常只有一项或少数几项能够帮助当事人实现协作。请注意问题 5 中当事人提出的第一个条件，尽管对 R 十分有利，双方之间的激烈争吵也只是修辞问题。“如果不是 R,是什么？”尽管当事方都希望在提出第一个条件后竭力回避字母 R,但是回答这个问题远不如多给几个盲目选择的机会那样容易。为了强调这一点的重要性，我们可以假设，裁判人员考虑到第一个条件已经破坏了整个游戏规则，决定有必要向选手宣布完全改变整个奖励规则。A 将得到 B 在问题 5 中的奖项，反之亦然。那么，第一个提出条件 R 的选手是否有理由改变选择？或者假设裁判宣布，无论他们选择哪个字母，只要双方选择的字母相同，则奖项保持不变。如果情况如此，他们仍然会对字母 R 趋之若鹜，因为这是他们有可能实现合作的惟一途径。如果我们回到游戏的开始，假设当事人还没有做出选择 R 的决定，我们可以想像墙上有个标识似乎在提醒他们：“在犹豫不决的时候，一定要选择 R。每个选手都能够看到这个标识，它仍是实现协作的一种方式。”这里，我们再回到那对夫妇在百货商店

走失的场景。假设夫妇二人都能看到一个醒目的标识，并想像其在对自己说：“管理人员建议，请走失的顾客到位于一楼角落的服务台等候。”如果这样的话，问题早就解决了。很显然，就他们的醒目符号源以及与其他双方惟一希望具有醒目性的标识而言，乞丐绝不可能是选择的主体。

在问题 5 中，如果对手知道你的奖品组合，而你却不知道他的（例如在一些问卷出现的类似问题 5 这样的情况），那么一场荒诞不经的表演也就宣告结束了。鉴于你没有足够的条件预测对方可能做出的选择，甚至没有机会向对方献媚或主动选择一个“公平”的折中方案，所以，双方实现协作的基础只能是你充分地解读自己所掌握的信息。也许你喜欢的字母也正是对方所喜欢的。似乎很难解释为什么选择这个字母，而不选择另外一个字母，因为你没有任何充足的理由可以说服自己：除了字母 R 以外，还有哪个字母对他更有利。对方对你的了解、你对对方的无知以及缺乏协作的替代条件，都会轻易地迫使对方做出对你有利的选择。（实际上，在小规模问卷调查所得到的众多结论中，这是一个非常重要的结论。）诸如一个空降兵知道对方的位置，而对方不知道自己的位置等类似案例也属于这种情况。[1]

显式谈判

在这里，为默式谈判定义的“协作”概念似乎还不能被直接应用于显式谈判。在谈判主体能够进行有效沟通的情况下，尚不需要他们出自本能的协作行为。提示符号的偶然出现，不仅能够帮助主体进行协作，而且能够影响在缺乏有效沟通条件下取得的谈判结果，还可以还原附带的各种具体因素的原有特征。

甚至在显式谈判中，有许多相关案例可以有效证明这一影响的存在。例如，在讨价还价过程中，人们更倾向于使用简单的数学表达形式。一个细节之处是人们往往喜欢使用整位数字（round numbers）表达结果。例如，如果汽车销售商能够将其产品的最低价格 2 507.63 美元中

[1] 在前面，还有一个脚注案例证明，谈判中的弱势有时可转化为优势。

的 7.63 美元去掉，则会吸引更多消费者的眼球。“折中”（split the difference）的建议通常能够促使谈判取得成功也证明这一点。当然，妥协的部分大多是一些细枝末节、微乎其微的次要部分。更加能够引起人们注意的是，诸如复杂的量化规则，特别是成本、利润分配等问题的旷日持久的谈判，往往最终都以一些简单的解决方式结束，如简单地平均分配，或参照国民生产总值、人口、外汇逆差等常量按比例分配，或按照双方在以前谈判中达成的比例，尽管有时可能在逻辑上毫无联系。[1]

先例所产生的影响也许远远大于逻辑的重要性和法律效力。一个成功解决罢工事件或国际债务问题的先验往往会在以后的谈判中被谈判人员不假思索地原本引用。当然，有些时候也需要制定统一的标准或类似情况得到类似结果，但是更多时候是，在一些引人关注的重大先例的阴影笼罩下，谈判情景中频繁出现冷酷无情的场面。[2] 同样的道理，调解人通常拥有决定是否达成协议或决定协议内容的权力。他们的解决方案之所以能够被人们所接受，很大程度上不是因为方案多么公正合理，而是由于当事人的接受和认可。“实地调查”(fact-finding) 报告也许能够通过提出可行性方案消除不确定因素的存在，从而将人们的视线集中到问题的焦点。但是即使如此，实地调查报告内容本身并非事实，它只是调查人提出的可能产生某种影响的具体建议而已。

同样，人们总是乐于安守现状或接受自然形成的界线，以至于一条条纬线最近也向人们展示了其作为实现协作关键点的强大生命力。不管怎么说，人们之所以将小河作为他们双方部队停火的分界线或仍然使用原有的分界线，自然有其合理的依据，无论他们现在能够发挥多大的作用。但是，与双方当事人达成明确协议的权力相比，地理条件就其自身的实际意义来说要逊色许多。

如果这些研究发现只是为了证明，谈判可以以某种简单有效的条件

[1] 在向联合国善后救济总署提交的众多建议案中，只有国民生产总值的 1%这一方案被接纳，其主要原因是计算公式间接，易于操作。可以肯定地说，美国代表在谈判中也倾向于这一方案。但是这个事实对本案例的积极作用和消极作用不分上下。

[2] 二战后，中东国家有关石油开采权的分配广泛采取五五分成的方案。这一案例有力地证明了本文的相关论述。

结束或使当事人之间寥寥无几的数美分或数公里路差距可以被轻易忽视，那么它们的价值可谓微乎其微。但是令我们感到欣慰的是，这些研究发现的作用不仅如此。这些研究发现揭示，谈判结果不仅反映了谈判主体间的权力平衡问题，而且将谈判的主动权从一方手中转移到另一方手中。愤世嫉俗者似乎能够根据双方争论的一些“焦点”问题，准确地预测出谈判的最终结果和现实形势所提供的条件，甚至根本不用考虑谈判中的有利条件，更不用说双方即将开始的争论和谈判中的压抑气氛。尽管没有现成可供借鉴的其他规律，但双方能够妥协的“显著”位置依然有规可循。如果谈判的“自然”结果反映了谈判双方的谈判技巧，那么，有必要把当事人的这种谈判技巧看做是他们向人们展示对自己有利的某种潜在结果的能力。谈判结果的公平性和公正性，也许并非像谈判双方之间的实力对比那样简单明了。

如果谈判形势已经决定了谈判的最终结果或达成妥协的关键点非常明显，这一结论也许看起来缩小了谈判技巧的生存空间。但是，它毕竟对谈判技巧发挥作用的时机产生了影响。在很大程度上，“明确”的谈判结果取决于谈判问题的自身特点或谈判双方对类似或相关谈判先例的理解，或谈判双方对谈判所涉及的相关信息的掌握程度。倘若委员会决定就成本的分担问题进行谈判，那么，影响谈判结果的因素将是多方面的。例如，是否分摊应付款和税费问题、后勤保障委员会是否充分准备所需要的有关国民收入和收支平衡的相关资料、谈判人员是否准备了此前亲身经历的谈判先例、将两个不同的问题按照同一议程进行谈判是否会产生重大影响并涉及两个问题之间的某些共同特征。当正式谈判开始，双方开始了一系列斗智斗勇行为。[1]

这些结论如果正如作者所认为的那样正确，我们对默式谈判进行的相关分析都有助于读者理解无时不在发挥作用的各种影响因素。也许默式谈判自身的逻辑性也能增强其说服力和可信度。默式谈判的根本问题

[1] 或许这一普遍方法隐含技巧的重要作用。如果一方无法将问题模式化，从而将问题的答案无限接近自己的立场，那么他可以继续将问题模糊化，寻找术语的多种解释并为模式化的信号增加“噪音”。这一策略也可能失败，但是在我们上面提到的所得税问题中绝对适用。

是协作问题。在以后的篇章里，我们还将进一步探讨影响显式谈判主体进行协作的因素。答案可能是，显式谈判要求谈判主体为了实现最终合作，必须具有某种协作的意愿。详细论述如下。

在大多数显式谈判中，只要谈判主体希望通过双方的妥协让步达成某种共识，谈判就可能存在多种结果。在这种情况下，至少有一种结果是一方或者双方希望通过妥协让步实现的，而且另外一方对此也十分清楚。同样，至少也存在一种潜在结果是一方希望按照自己的立场达成，即使他也没有充足的理由说服自己。因为对方能够肯定或者预测双方宁肯妥协让步，也不愿空手而归。谈判双方都是根据自己对对方立场和意图的预期判断，来确定自己的立场和谈判策略。谈判双方心里都清楚，对方也会有与自己一样的行为，最终的谈判结果也体现了一方对另一方的让步底线。谈判主体预期的主要构成是，一方对另一方预测自己对其预测的预测。因为除此之外，在这种复杂多变的谈判形势下，谈判方似乎别无选择，双方只能以这种方式做出自己的决定。这一无限循环的推理模式最终都将回到一点，即一方预测对方已经退到了让步的底线。

这时，在我们的脑海中出现一个疑问：是什么因素导致双方的这种无限循环推理模式和整个谈判过程的结束？我们推测，可能是双方的某种直觉偏好在发挥作用，特别是对明朗、独特、简单、先验的偏好，或者某种能够促使谈判主体将谈判结果与其他可能性相区分的潜在规律。我们不禁反问，既然双方预期目标与实际达成的谈判结果只是程度上的不同，为什么谈判双方还要想方设法向对方展示自己信誓旦旦、宁死不屈的坚定姿态？答案可能是，双方必须为各自的坚定立场寻求充分依据，并且在类似的立场观点中找到它们之间的不同“规律”。这一规律也许并没有客观存在的“关键点”那样令人可信、坚不可摧，但至少在谈判中能为各自辩护：“不是这个结果，还能是什么结果？”

但是这些因素还不足以促使谈判双方达成共识。如果一方打算妥协让步，那么他必须控制对方对自己的预期判断，知道自己妥协让步的底线。如果一方不希望自己的妥协行为变成自己软弱的表现，他必须在谈判过程中一个非常醒目的点上停止让步。这个点可能来自于调节人的提示或任何新旧条件的连接点。如果一方的要求就是60%，但是能够接受

的条件是 50% ，那么一旦他接受 49%这一条件，对方就可能认为他已经到了崩溃的边缘，进而将提出更高的条件。

现在，让我们再次回到前面的两军抢地案例。如果一方撤退到小河的沿岸，这一行为可能会被对方看做是防守待攻、立场坚定的信号。但是，倘若该部队继续退让到小河的背后，这时另一方就可能把这一行为当做对方已经溃不成军、立场动摇的信号。同样，进攻的一方如能不战而屈人之兵，将对方逼迫到小河沿岸，他就能迫使对方接受自己苛刻的条件，双方可能在小河沿岸或其他什么地方保持对峙。

在作者看来，从直觉的角度讲，这一假设能够解释任何条件下关键点的合理性问题。但是，如果这一假设不能对默式谈判进行具体的逻辑解释，在旁观者看来还是含糊不清、难以理解。对默式谈判的逻辑解释不仅能够提供逻辑推理依据，而且还能揭示各种相关心理现象——如双方为实现预期目标的心理协作——存在的可能性，甚至在一些教科书中具有非常大的说服力。在缺乏沟通的情况下，双方的心理协作如同行为的协作一样重要。事实上，二者都或多或少地涉及双方对对方进行本能的预期判断。因此，一些有关默式谈判博弈的结论和协作产生共识的逻辑作用经过实证检验后证实：双方能够通过协作达成共识，以及当协作共识成为必要条件时，形势的客观条件能够产生绝对性影响。在缺乏沟通的情况下，双方能够产生“某种”共同需求。毫无疑问，尽管在沟通成为可能时，这种需求非常微弱，但一定存在，因为沟通的存在并没有削弱 5–5 分配方案的对称性、小河的独特性和 ABC 的自然排序规则。

也许我们探讨的所有问题的根源来自于默式谈判的自身逻辑。与心理因素在显式谈判中的突出作用一样，默式谈判的逻辑也许是最广泛的推理问题。同样，倘若我们希望归纳结论来自于对实际谈判中种种可行性结果的研究分析，我们也许不愿意承认偶然因素的作用。但是，有关默式谈判和显式谈判的案例已经强有力地证明，在二者之间完全有可能存在某种共同性或者相似性。

为了说明这一点，我们这里回到前面提到的平分 100 美元的案例中。现在看来，对 100 美元五五分成似乎具有可行性，但是产生这种可行性的原因是多方面的。也许这种方式非常“公平”；或许体现了权力

平衡；或许正如这上面所论述的，这种方式能够向双方传递这么一个信息，即：双方都乐于接受这种方式。我们对默式谈判的研究分析为后一种观点提供了依据。这个依据是，如果双方在缺乏沟通的条件下分配100 美元，他们得到的分配结果可能是 50-50 。这并非出于直觉。我们认为在略有不同的语境中——默式谈判语境——我们的结论客观上具有解释力。

两个指挥官的案例再次说明了这一点。该案例中，如果双方的生存一定程度上取决于他们设定的脱离接触线和有效沟通的缺乏，他们可能会把小河当做他们之间默式谈判的关键点。这验证了双方指挥官为什么承诺小河为双方的停火线或没有承认小河为停火线。所以，默式类比方法至少能够证明，“协调预期判断”（coordinating expectations）的思想具有积极意义，而非神秘莫测（mystical）。

也许我们能够进一步阐述这一观点。倘若双方都乐于接受这一标准衡量，那么，谈判结果的惟一显著特点就是非常“公平”（fairness）。我们可以认为，双方关注的“公平”结果增强了公平的道德约束力，如果这能够填补不确定因素留下的空白的话；同样，我们认为公众舆论要求谈判双方达成的结果必须“公平”或“合理”。我们或许夸大了公众舆论的作用或至少误解了公众舆论对主体施压的方式，除非我们相信公众舆论能够协调谈判双方的预测判断。换句话说，暗示（suggestion）以我们这里提到的方式促使公众舆论、先例或道德标准产生有效影响。为了再次证明这一结论，我们只能认为谈判主体必须在缺乏沟通的条件下达成最终结果，并且将舆论和某种道德规范约束看做是类似于前面案例中的暗示。问题 7 中的调解人也发挥了类似的作用。总之，尽管促使公众舆论影响谈判双方行为的因素看起来可能是道德的责任性和敏感性而非某种暗示，但是我们仍然有必要关注公众舆论背后的动机。作者认为，对简单有效的基本规律的渴求反映了我们这里讨论的机制问题。

但是，如果我们的推理主线有效，任何对显式谈判的研究分析必须关注我们提到的谈判沟通问题，以及谈判双方如何释义现实条件提供的各种枯燥无味的细节。这意味着默式谈判和显式谈判并不是毫无联系的两个概念，而是从默式谈判的部分、错误和有限沟通逐渐过渡到显式谈

判中谈判双方视情况而定的完全沟通。因此，二者在某种程度上都表明，谈判双方均受到某个预期目标的影响。

这一观点并不期望，明确可见的谈判结果限制缺乏沟通的谈判主体的行为。除了我们为了研究分析的需要而人为设计的案例，沟通存在时的谈判关键点与缺乏沟通时的谈判关键点绝对不同。尽管如此，默式谈判的某些重要原则显然至少也是显式谈判的重要原则之一。甚至众多所谓的“显式”谈判通常也包含许多非显式因素，如谈判双方的动机、间接沟通、立场的隐瞒、故作高深、被众多参与者和冲突利益所迷惑、对默契和暗示的需要等等。这些因素无不对双方是否能够产生默契具有极大的影响。

或许，许多社会稳定和利益集团形成的问题也同样受到地域或环境因素的影响。例如，寥寥数人的政治集会可能会发展成为规模庞大的盛会；当出现政治真空或无政府状态时，宪法的有效性则需要主体公众的支持；一个犯罪集团元老的权力可以影响整个黑社会，因为服从是建立在服从者惩罚违抗者这一简单的逻辑基础之上。我们通常所指的社会“号召力”也反映了同样的道理。价格领先、各种非价格竞争或价格稳定等经济现象无不证实默式沟通的重要性，及其依赖于谈判主体对现实情况的准确解读。不谋而合的反叛行为告诉我们：当领导者不堪一击时，人们通常需要一种默契行动的暗示。这种暗示使他们相信，如果自己起来造反，对其他人也会做出同样行为，这样就形成了不谋而合的反叛行动。（这种暗示也可能来自外部甚至是反叛主体，因为具有领导潜质的反叛者本身也具有某种号召力。）

默式谈判和有限战争

我们进行的研究活动是否能够帮助我们解决实际中遇到的默式谈判问题，特别是战略行为或有限战争问题？可以肯定的是，即使没有显式谈判，我们也一定能够找到产生有限战争——真实战争、争夺管辖权的战争等等——的有限条件。但是，我们也知道其可能性非常之渺茫。朝鲜战争是一场有限战争，第二次世界大战中也没有出现毒气战。与前面我们讨论过的案例相比，这两个历史事件足以证明有限战争的可能性。

假设研究分析能够得出有效的结论，那么这个结论一定是关于如何有效地寻找谈判的成功条件，而非探索默式协作成功的可能性。

如果一定要得出研究结论的话，这些结论无疑是：（1）无论是默式谈判结果还是通过不公平或艰难谈判取得的结果，都需要比其他淘汰结果具有质的优越性，而不仅仅是程度的问题。（2）如果非要在沟通匮乏的条件下实现结果，谈判双方应该具有足够的宽容心，允许现实条件影响谈判结果；只要是双方的共识，允许谈判结果对一方不利或者没有任何意义。

二战没有出现毒气战。这一协作也许并不是战争双方战前的约定，而很大程度上是战争双方的默契行为。因为很难想像，战争双方在无法（或能够）进行有效沟通的条件下，还能就有关毒气使用达成类似共识？当“有些”毒气的使用涉及的不是是否使用而是使用程度的问题时，“禁止使用毒气”是一种最简单、最有效的方式。尽管诸如毒气使用对象仅限于作战人员、禁止将毒气应用于进攻行为、使用可测毒气前必须通知对方等规定中的某些规定确实能够限制战争双方的行为，但是“禁止使用毒气”可以简单地使其成为双方惟一合作点，因为任何一方都无法找到其中的漏洞，对方也不用为此担心。

朝鲜半岛独特的地理特征有助于分析战争的有限性和地缘的局限条件。朝鲜半岛三面环海，北面边界鸭绿江清晰可见。三八线似乎毫无疑问地成为双方对峙的界线。三八线成为战争双方的最优选择，其原因并不是朝鲜半岛的这个“腰”战线较短、易于防守，而是据守该点不容易被对方认为是进攻（退让）举动。

台湾海峡的天然屏障使中国大陆共产党政权和台湾国民党政权之间的僵持对立成为可能。这并不仅仅因为台湾海峡增加了跨海作战的难度，而是因为台湾岛是一个统一整体，四周的水域成为其天然屏障。如果牺牲台湾岛的任何一片区域，都会造成当前台海局势出现不稳定因素。对中国大陆而言，情况亦如此。双方在台海两岸的任何举动都是一个程度问题。而如果有一方决定跨海作战，那就宣告这种“默契”的结束。

在朝鲜战争中、核武器与常规武器的差异限制了武器的使用。很难

想像，如果美国使用原子弹对北朝鲜进行选择式目标打击，战争双方还能达成类似的默契。[1]没有任何涉及规模和目标的定义像“零规模，零目标”这样清晰易懂、不言自明。美国对深陷印度支那战争的法国提供援助也仅仅局限于物资援助，而非人员援助。当然，法国希望美国能够给予空中支援，哪怕是有限的空中支援。但是，殊不知射出去的箭难以收回。一旦美国答应了法国的请求，给予有限的空中支援或地面支援，那将意味着更大规模的干预。一方只有采取全部撤军行动才能表明其真正的撤军意图，很难保证采取部分行动或将规模控制在有限范围之内。

战争双方对涉及有限条件的沟通或协作的需要，极大地影响了报复战略。地区性侵略局限于一个区域。如果幸运地拥有一个天然屏障作为边界，这也许将迫使对方默认其地缘劣势和进攻目标的局限性。一方和双方也许会乐于接受部分失败，而不敢冒天下之大不韪首先破坏双方达成的默契以导致对方采取相同的举动。这一默契能够得到双方的尊重，而且一旦被打破，谁也无法保证能够达成新的默契，并就遏制冲突扩大及时达成共识。但是，如果报复方能够掌握报复的方式和地点，那么受害者将很难得到报复者提出的有限条件，结果造成自己不得不在其反报复行动中接受对方提出的条件。事实上，一方首先发动报复行动，就如同某种独立宣言，告诉对方自己不会与其达成任何共识。因此，如果侵犯者自身行为隐含的寓意巨大而无法让人接受的话，那么，双方就战争有限性条件达成共识就变得异常困难。

综上所述，有限战争问题并非涉及主动权在双方之间周而复始的转移。这是一个错综复杂、纷乱不堪的世界。这个世界更承认质而非量的差异，不仅存在众多选择机会，而且迫使双方接受各种因素的制约。作者认为，凡是具备上述条件的有限竞争都属于这一类情况。

前期准备

尽管本文论述的默式谈判不仅具有可行性，而且经受了系统分析的检验，但是我们还是无法保证默契在任何条件下都能发挥作用。或者即

[1] 有关问题在附录 A 有详细论述。

便如此，我们也无法保证默式谈判产生的结果优于完全沟通条件下所产生的结果。当然，谁也无法保证下一次战争爆发时（如果可能的话），战争各方能够及时找到有限战争要求的有限性或者及时找到一种有效的自我保护方式，除非显式谈判能够出现。鉴于此，我们有必要考虑在默式谈判开始之前进行必要的准备活动，以增大谈判成功的几率。

很显然，保持沟通渠道的畅通具有很重要的意义（至少应该允许投降信息的互动传递）。这一原则的技术层面涉及发收信息的主体、相关的部门机构、传递的中质（如果存在的话）、传递的备用方案（假设中质和设施被人为破坏）等因素。作为防止核战争的努力，战争双方只有一次机会判断有限战争是否已经发展到了极限和一场全面战争已经开始，因为在 24 小时内还无法沟通就可能丧失最后一次保持战争有限性的机会。

有人认为，有必要发挥中间人或调节人的积极作用。一个能够发挥重要作用的调节人必须事前得到战争双方的首肯，至少存在一个友好的先例、传统或迹象。甚至即使我们排除了为类似紧急事态所做的相关准备工作，战争双方如果有调解的意愿或者哪怕做出微弱的努力，都有助于我们在形势极度恶化的紧急事态中，建构一个最有价值的调解机制。

但是如果敌对双方缺乏和解的诚意，准备阶段的各种努力都将付诸东流。如一方拒绝做出任何友好的姿态或者身入潜在战争的一方甚至希望在一场两败俱伤的战争中赢得战术上的利益。你也许会问，为什么会这样？原因很简单，威胁、欺骗和威慑战略在从中作梗。国家领导人对有限战争的极度乐观加速了战争的爆发，或可能导致战争的种种举动，不管其最初动机是出于侵略还是出于报复。具体来说，美国是否决定以核进攻报复地区性的侵犯行为，很大程度上取决于——苏联人知道这一条件——我们对报复行动自身有限性的预期判断。换句话说，当双方迫切需要确定有限战争的必要条件时（假设双方愿意输掉这场有限战争），这取决于我们如何预期判断，美国人和苏联人能否找到有限战争的必要条件并达成默契。假设苏联人不愿意采取任何可能导致有限战争的举动，我们将放弃实施打击报复。但是，他们也许会为化解我们打击报复的威胁敢于突破这些有限条件。这里，我们再次回到前面提到的两个空

降兵的案例。其中一个空降兵明白，如果对方在跳伞之前知道他们会合后会得救，那么，对方也就不会关心飞机的安全问题。因此，只要一方拒绝就紧急情况与对方进行沟通，则对方就会像热锅上的蚂蚁，担心自己将不得不在陌生的地形中苦苦探索，寻求自救。

尽管对方的这种考虑或对重要谈判不利的种种消极因素削弱了双方前期协调的可能性，但是前面的博弈游戏给了我们一个启发，即以协调预期判断为目标的谈判或沟通不一定需要双方的参与：单方面的谈判也可以实现双方的协作，从而实现自救，即使一方人为地制造信息沟通的障碍。你是否还记得其中一个案例中，那个建议对方选择字母 R 的一方？只要对方明白——他当然明白——字母 R 是惟一没有风险的建议，很显然双方一定会达成共识。（即使对方拒绝接受字母 R 的建议，只要没有歧义的反对意见出现，这就只能说明对方意识到了这一点，并不能否定该建议的积极作用。）同样的道理，如果在空降兵怎么也没有想到飞机会失事的情况下，有一个空降兵调侃：“如果我跳下去一定会遇到敌人，我就首先跳到能看到的第一个山头上。”当飞机果真失事时，另一个空降兵也许会回味前面那个人的话，而且也会那么做，尽管他此前差一点嘲笑前面那个人“真愚蠢”或“我绝不那么做，攀援会弄伤我的腿”。假设双方迫切需要某种达成共识的暗示并且发现了这个暗示，那么，不管该暗示多么微弱或不公正，一方在另一方缺席的情况下一定会接受这个暗示。因为一旦发生波及双方利益的突发事件，二者将为了迅速地达成某种共识而暂时搁置双方的利益冲突。

第二部分

博弈论的重新定位

第 4 章　趋向互动决策理论
第 5 章　实施、沟通与战略行为
第 6 章　博弈论与实验研究

第4章

趋向互动决策理论

在完全冲突的战略（零和博弈）研究领域，博弈理论表现出了其独特的洞察力，并取得了丰富的研究成果，从而为该领域的发展做出了巨大贡献。但是与此形成鲜明对比的是，在冲突与共同利益并存的行为战略（非零和博弈）研究方面——诸如战争及战争威胁、罢工、谈判、预防犯罪、阶级矛盾、种族冲突、价格战和黑邮件；官场中的勾心斗角或交通堵塞中的你争我抢；以及对孩子的管教——传统博弈理论的表现相对而言要逊色许多。在这些“博弈游戏”中，尽管冲突涉及许多重大利益，互动却是逻辑结构不可或缺的一部分，并且要求冲突双方某种程度上的合作和互谅——不言而喻的默契——即便只是为了避免两败俱伤。在一些博弈游戏中，尽管保密工作可能发挥着战略性的作用，但是有时也需要谈判博弈双方进行必要的交流和沟通。最后，在这些博弈中，博弈双方都能够采取有效行为以避免两败俱伤的局面发生。鉴于此，拥有主动权、人才优势和自由选择权的一方在实际博弈中并不总是处于有利的地位。

就大多数情况而言，传统博弈理论通常将完全冲突的战略研究证实了的方法和概念应用于互动博弈场景（非零和博弈）。在本章和下一章中，作者通过把“零和博弈”看做是具有有限性的场景 (limiting case) 而非一个分切点（departure point) 而拓展了博弈理论的范畴空间。作者

对博弈理论的假设部分伴随两条主线：一条是旨在研究构成博弈双方预期判断的直觉因素和启发性因素；另一条（在下一章中将详细论述）是研究实际战略博弈中存在的基本行为和促使这些行为产生的结构性因素。后者将涉及一条基本概念，如“威胁”、“实施”和沟通能力或破坏沟通能力。

发展中的博弈理论及其两条主线反映了二者对“零和博弈”的偏好。在已经被人们广泛接受的“零和博弈”理论中，建议与干预、威胁与允诺都无法产生积极的结果。其中主要原因在于它们都主张，博弈双方之间的关系要么完全友好，要么极端不平衡。这就可能导致不利一方采取最小最大战略破坏这种不平衡关系。如果可能的话，这种行为甚至建立在随机化机制的基础之上。因此，博弈双方在非敌即友的完全冲突场景中追求的“理性战略”不可能实现双方间的互谅或为一方利益而存在的互动。

假设“零和博弈”是一个具有有限性的完全冲突场景，那么与之相对应的另一个极端又是什么呢？答案一定是“完全合作”(pure collaboration)博弈。在后一种博弈场景中，双方的利害关系相互依存，二者之间存在达成某种共识的基础。不管各自是否赢得双方共同利益中应得的那部分还是额外的一部分，他们都必须按照自己的利益需求，排列出所有可能的结果。（或者为了在开始就取得共识，双方应该清楚他们的利益必须一致，因此双方之间的信息沟通中不存在利益的冲突。）

那么，完全合作场景与博弈理论或谈判之间存在什么必然联系呢？一个略显偏颇的答案——只是为了在此强调博弈场景一无是处——认为完全合作博弈存在非零和博弈中经常涉及的沟通和认知问题。无论沟通结构何时难以允许双方根据预定计划分配他们之间的任务，双方在博弈游戏中都很难协调彼此的行为。为了协调彼此行为，双方需要了解对方的情况，研究对方的行为模式；为了建立共同的行为模式和信息释义系统，双方还需要反复不断地沟通协调，沟通的方式也许是某种暗示或默契行为。诸如努力避免相撞的两辆汽车、共跳陌生舞曲的一对舞伴，以及战斗中失散的游击队员都需要类似的默契，这类似于音乐大厅的听众不约而同地鼓掌致谢。

如果国际象棋是标准的零和博弈游戏，那么，看手势猜字谜的游戏就是典型的完全合作博弈；如果追击现象形象地描述了零和博弈，那么，集合现象则对合作博弈具有同样的意义。

O.K 穆尔（'O.K Moore）和 M.I.伯科维兹（M.I.Berkowitz） 做的一个实验证明了两个有限博弈场景存在的可能性。[1]这个实验属于零和博弈实验，共有 6个人参加，每组 3人。这 3个人具有相同的价值观，但是由于实验的特殊性，每组不是一个整体。实验的特殊性在于，每组成员必须彼此分开，只能通过 6部电话沟通。而且 6部电话共享一条电话线，每个人既能听到自己同伴的讲话，也能听到对方成员的讲话。但是，同组成员之间不准事先约定任何暗号。这样，在两个小组之间就出现了完全冲突博弈场景；每个小组的成员则出现了完全合作博弈场景。

假设在这个充满技巧和机会的博弈试验中，我们压制“对方小组”，并使该小组成员在缺乏沟通的情况下试图合作制定一个取得胜利的战略，那么，我们就使该小组成员之间形成了完全合作博弈关系。研究人员研究分析了实验或现实中许多类似的“博弈游戏”。事实证明，非零和博弈与组织论或信息论存在很大共性。 [2]

第 3 章中的实验表明，即使在缺乏有效沟通的情况下，谈判双方完全可以通过协商做出共同的抉择。而且，这些实验还表明，在默式谈判中，谈判双方之间的共同利益完全可以克服二者之间的利益分歧。在这些情况下，完全合作博弈体现了与非零和博弈的根本不同之处。

在这个以谈判双方相互沟通为基础、旨在研究协作解决问题的实验中，我们发现一个充分反映非零和博弈特征的现象，而且与“零和博

[1] O.K Moore and M.I.Berkowitz,Game Theory and Social Interaction,Office of Naval Research, Technical Report, Contract No.SAR/NONR-609 (16) (New Haven, November, 1956) .

[2] 对协作问题的相关深入研究，请参阅 Jacob Marschak “Elements for a Theory of Teams”, and “Toward an Economic Theory of Organization and Information”, (*Cowles Foundation Discussion Papers*) ,Nos.94 and 95, (New Series)。相关的文章还有 Jacob Marschak and Roy Radner) , “Structural and Operational Communication Problems in Teams” ,*Cowles Foundation Discussion Papers*,Nos.2076.Alex Bavelas, “Communication Patterns of Task-oriented Groups,” D.Cartwright。

弈”存在几乎相同的关系。虽然都是冲突合作博弈，但是一个轻视合作，另一个是轻视冲突；一个重视保密，一个重视沟通。

需要强调的是，从严格的技术意义上讲，完全合作博弈是一种战略博弈。这种行为模式强调，一方的最优行为选择取决于其对对方的行为预期判断，反之亦然。这种互动行为预期判断将战略博弈与机会博弈和技巧性博弈区分开。在完全合作博弈中，谈判双方利益共存；而在完全冲突博弈中，谈判双方利益冲突。但是无论是哪一种情况，如果谈判双方之间不存在互动依存的关系，那么二者都将不可能做出正确的行为选择。[1]

现在我们回到福尔摩斯 (Holmes）和莫日娅蒂（Mariarty）的精典案例中。二者分别乘坐不同的火车，并且无法进行沟通，但是两人必须决定下一站是否下车。现在我们可以考虑三种不同的结果。第一种情况是双方在不同的车站下车，那么福尔摩斯将得到奖金；反之，莫日娅蒂将得到奖金。这是一种零和博弈，双方的利益完全冲突。第二种情况是双方在同一个车站下车，不论双方在哪一个车站下车，只要二者在同一个车站下车，那么两人将同时得到奖金。这是一种完全合作博弈，双方的利益互存，一方之所得取决于另一方的合作。第三种情况是只有在同一个车站下车，双方才能获得奖金。但与第二种情况不同的是，如果两人在某一个特定的车站同时下车，那么福尔摩斯得到的奖金多于莫日娅蒂；而如果两人在另一个特定的车站同时下车，那么莫日娅蒂得到的奖金多于福尔摩斯。如果同时在不同的车站下车，两人都将一无所获。这是典型的非零和博弈，或者说“非极同选择” (Imperfect-correlation-

[1] 关于这一点，参见 Carl Kaysen 给 Von Neumann 和 Morgenstern 的文章 “Theory of Games and Economic Behaviors” ，他在评论中写道：“战略博弈理论仅仅涉及几个人的互动行为。该理论认为，一般情况下，根本不可能量化博弈双方之间的行为互动。实际上，这个理论根本不存在所谓的量化。”R.Duncan 和 Howard Raiffa 在其著作 *Games and Decisions* 中作了相关论述：“表面上看，利益冲突问题对双方而言是风险与动荡并存的条件下的个人决策问题。不稳定性和动荡来自于一方对另一方行为的无知。”（第 14 页）从上面的论述，我们可以看出他们的观点更倾向于冲突，合作因素在他们看来微不足道（第 59、88 页），他们往往把每个参加者看做独立个体（第 13 页）。

preferences) 博弈。在这种博弈场景中，博弈双方的冲突与共同利益并存。为了强调沟通和信息系统对双方的重要性，我们可以将实验进一步细分成使 1 或 3 个变量对一方更有利。

上面游戏中的三种情况揭示了战略论的基本要素：双方的对应选择取决于二者之间的互动预期判断。两人需要不断地对对方的意图进行预测，并不断协调双方的行为，从而实现双方的利益最大化。

博弈再分类

在进行深入讨论之前，我们有必要对相关的博弈场景进行再分类。人们通常将博弈简单地分为零和博弈和非零和博弈两种情形。但是，这种分法缺少我们需要的对称性，并无法将有限博弈与零和博弈有机地区分开。我们的分类要点是，一个二维图表完全能够解释双人博弈。对双方而言，每一种博弈结果的收益可以由图表中某个点的两个坐标值表示。完全冲突博弈可能得到的所有结果由图表中一条斜线上的部分或全部点表示。在混合博弈或谈判中，至少存在两个点构成一条线代表完全合作博弈。[1]

这里，我们可以使用惯用的博弈术语，可以称完全博弈为定量

[1] 如果博弈的本质就是为了帮助选手使用随机手段做出战略抉择，或借用机会机制通过谈判达成有效的协议，那么将存在通过合作达成协议的空间，甚至不是最好的结果。那样的话，代表完全冲突博弈的坐标一定会与水平线相交，其中两个轴点将代表博弈论中的选手“效用”。这一限制条件同样适应于共同利益博弈中。在共同利益博弈中，选手双方的选择排序相同，利益完全按照五五分成原则。完全冲突博弈和共同利益博弈要求博弈双方要么完全冲突要么完全利益共存，与此同时，二者表明所有相关战略的预期价值。代表其价值的两条坐标线的两轴分别代表选手的效用。这也验证了，代表结果的点必须处于水平线上。

同时，完全博弈不允许“单方面的补偿”。假设博弈双方间能够进行有效的沟通，一方以破坏将要达成的结果威胁另一方，要求其支付相关费用，这就出现了双方利益的冲突。实际上，代表贿赂款的坐标点总是或高于或低于文中提到的某个线，重新建构混合博弈的分配结构。如果完全冲突博弈的一方以破坏相威胁或以补偿相引诱对方妥协让步，那么就可能出现谈判的可能。二者之间的关系也不再是简单的完全冲突关系，代表受到破坏的损失或补偿价值的坐标点将脱离下斜线。换句话说，博弈应该允许产生一切可能的结果。（两个同时进行的完全冲突博弈即使受到水平线的限制，它们也为谈判留下了空间，除非两条斜线重合。）

(*fixed-sum*）或定比例 (*fixed-proportions*）博弈；对于有限情况以外的博弈，我们可以称其为变量博弈或非定比例（*variable-sum-variable-proportions*）博弈。鉴于博弈结果与双方的共同合作存在很大关系，我们也可称其为极负关系（perfect-negative-correlation）博弈和极正关系（perfect-positive-correlation）博弈，其他范围更广的混合博弈则可称为“非极关系”(Imperfect-correlation) 博弈。

要为冲突与互存同在的混合博弈取个恰当的名字并非易事。更有趣的是，我们找不到一个准确的词来形容博弈双方之间的关系：在合作博弈中，我们倾向于称双方为“伙伴”；在完全冲突博弈中，我们称二者为“对手”或“敌人”；但是在涉及战争、罢工、谈判等行为的混合博弈中，我们又找不到一个完美的词来概含其广泛的内容。[1] 在本书以后章节中，我将称混合博弈为“讨价还价博弈”(*bargaining game*）或“混合动机博弈”(*mixed-motive game*)，因为这两个词似乎准确地表达了混合博弈的精髓。当然，混合博弈并不是指博弈双方选择的混合性，而是形容双方之间关系的复杂、难以琢磨——互存关系与冲突关系或伙伴关系与竞争关系的混合。非零和博弈包括混合博弈和完全合作博弈。协作博弈似乎更能恰当地形容利益的完美结合，更准确地描述相关问题和行为。

协作博弈

虽然本书大部分内容都涉及混合博弈问题，我们还是有必要在此探讨完全协作博弈问题。这不仅有助于我们理解完全协作博弈的重要价值，而且将帮助我们辨析完全协作在有限条件下所表现出的某些混合博弈特征。

第 3 章中提到的各种完全协作博弈场景都不同角度地涉及了协同决

[1] 有必要强调，无论是在冲突论还是在伙伴论中，非零和博弈都能被恰当分类。为了更好地说明有限战争等问题，最好能使用准确表达双方共同利益或军事行动所涉及的谈判问题的相关词语。正如第 9 章将要谈到的，突袭问题本质上也是伙伴原则问题。假如博弈论过多地注重冲突问题，那么诸如互动决策将成为一个中性词，即包含两个有限博弈和混合博弈问题。

策的某些焦点，实现合作的暗示条件以及实现双方共同预期的一些规律。有人认为，同样的合作暗示条件不仅存在于完全协作博弈中，而且也存在于内含冲突的混合博弈场景中。事实上，有关实验证明，在缺乏有效沟通的条件下，上面的结论一定成立。但是许多案例表明，完全协作博弈——寻找伙伴和实现协作的默契过程——本身是一个重要的现象。不约而同或一呼百应的反叛起义行为充分说明了这一点。

出现不约而同的反叛起义的前提条件是，希望参加反叛起义的潜在成员不仅一定要知道集合的地点和时间，而且还要知道统一行动的具体时间。当然，这个问题只有起义领袖能解决。但是起义领袖往往是当局为了镇压起义而仇视和消灭的首要对象。为了解决这个问题，在领袖不在的情况下，反叛人员往往寻找一致行为的共同标志或暗示，从而使每个参加起义的人相信如果自己这么做就不会孤立。某个偶然事件往往成为协调起义行为的标志，并充当领袖的化身，发挥着沟通中质的作用。假设没有偶然事件，起义人员往往很难实现步调一致的统一行为，因为起义的敏感性迫使每个人必须知道共同行动的时间。同样的道理，在一个缺乏“醒目”中心地点和标志性建筑的城市，起义很难在短时间内聚集起来。因为没有一个地点如此“醒目”，从而使每个起义人员都知道别人也会到这里。领袖大选或其他选举活动中的乐队花车一定程度上也得依靠相互认识的标识，尤其当人们希望加入游行队伍或至少观看游行活动时。[1]

对默契合作和行为的依赖或许会产生我们不愿意看到的过分偏激行为。当白人和黑人都认为某个区域必然为黑人所占有时，这里的“必然性”就是双方预期判断的共同点。[2] 让人首先感到必然的不是最终结果而是这种预期判断；反过来，这种必然性的预期判断也导致最终结果的必然性。大家都希望别人也期望别人能够预期判断出最终结果，没有人

[1] 有个非常相近的例子：一个学生为了躲避背诵课文试图逃进游行队伍，结果却被一帮家伙抓起或“选举”到一个大家都不愿意做的差事。

[2] 参见 M.Grodzins 的 *Metropolitan Segregation, Scientific American*, 197:33~41 (Octomber, 1957)。该文深入探讨了这个现象。在具有选举意味的默契沟通基础上达成的协同预期，如同引起人群狂笑的窃笑。第四共和国或巴蒂斯塔王国的覆灭就是一个典型。

能够抗拒这种行为。除非是处于两个极端，否则没有一个点永久不变。因此，没有人能够预期将在某个点达成协作，如 10%、30%、60%；也没有哪个百分点刻意要求达成协作或提供一个集合地点。如果通常惯例是 100%，那么，只有明确的共识才能打破这个惯例。如果双方必须通过某种默契才能实现合作，那么，妥协让步就无法实现。人们往往受到不完全沟通条件的制约，可能很容易“同意”（默契地）变动，或却很难“同意”（默契地）稳于一处。住宅小区或学校发展规划等方面的配额制度可以被看做是，以具有沟通条件的有效显式博弈取代要求默式“解决方法”的默式博弈的努力。

协作博弈也可能隐藏于制度和传统的稳定现象或者领袖现象的后面。在一大堆规则中或许就存在制约冲突的因素；传统则是每个人期望别人作为潜在候选人保持清醒的特定规则。当默许无法发挥作用时，传统因素通常可能产生重大的影响。社会礼节和习俗——包括（标点符号的使用）被权威和相关因素侵占的部分——的作用也取决于它们成为协作博弈的现成答案：每个人都期望别人遵守法律习俗，从而违背者将受到应有的惩罚。时装潮流和摩托车爱好者也反映了一个博弈游戏。在这些博弈游戏中，任何人都不甘落伍于形成中的潮流以及盲目地阻止这种潮流形成。社会学的“角色”（*role*）概念——自己对别人的行为预期判断或别人对自己的行为预期判断——一定程度上可能被释义为“共同预期”的稳定性表现，或是协作博弈中的同一类型。一个人总是陷入自己的角色或是别人的角色之中，因为这是默认条件下的惟一角色。

现实生活中存在很多这方面的例子，如军队或舰队的团体精神（*Esprit de corps*）、某个大学或社团的价值观等。虽然社会中很多有机体在不断演变，但是它们还是不同程度地保持某种特有的稳定性，而且新成员的加入并不足以削弱这种特质。单位成员的个性似乎很大程度上涉及共同预期判断——每个人对别人的预期判断——对新成员的价值观念的塑造有助于对以后成员价值观的塑造。这具有“社会契约”的味道，每一个新成员也会认识到并接受这一“契约”。有人告诉我，社会实体传统的连贯性是人为地保留某个军队编制或番号的主要依据，尽管该部队可能已经到了被淘汰的边缘。这个案例表明：与组织共存的价值观中

该组织未来发展不可或缺的必要条件，应该受到保护和尊重。同样的道理，在一些国家能够实施个人所得税政策，而在另外一些国家却无法实现。原因很简单：假设国民之间存在某种预期价值判断，那么，尽管每个人都打算在不触犯社会共识的情况下逃税，但结果是每个人可能出于对彼此负责或对法律惩罚的惧怕而去缴税。因而，大家都能够验证自己的预期判断。

协作意识的本质　在此我们有必要强调，协作过程并不是简单地对“普通人”（average man）的行为预测。在默式协作情况下，协作也不是仅仅对别人客观情景中的行为预测，而是别人试图预测自己如何预测对方的行为过程，反之亦然。（在报纸个人专栏中“遇见”某个人就是一个典型情况。[1]）除非客观现实能够提供某种共同选择的暗示条件，否

[1] 同样的情景还有，以人类能够接收的波段接收外星人的信息。“我们应该使用什么波段？长波搜集一个未知频率的微弱信号非常困难。但是，在最利于电波传播的地区存在一个惟一客观的频率标准，应该尽快将这一标准通知整个宇宙的每一位天文观察者：最佳波段为中氢 1420 兆。”［吉斯普·科科尼和菲利普·莫里森，自然，（Gueseppe Cocconi and Philip Morrison,*Nature*）,sept.19,1959,pp.844~846］约翰·利尔应用了这一理论：“地球上每一位天文学家都会说，为什么，因为 1420 兆是最佳中氢波段。氢是地球外层最丰富的气体。我们的邻居中甚至是最菜鸟的天文学家也能观察到。”［“寻找天外生命”（“The Search for Intelligent Life on Other Planets”,）*Saturday Review*,JAn.2,1960,pp.39~43］　寻找什么信号？科科尼和莫里森建议为一组脉波或简单的数字。

这表明了实验的另一个趋势。在一系列随机的红灯和绿灯后，实验客体需要猜测下一个出现的是红灯还是绿灯。他们一定会以以前的认知进行猜测，即“非理性”行为模式。正如赫伯特·西蒙（Herbert Simon）所言“人类这种动物不仅仅具备学习能力，还具有求知行为模式并建构概念。”（Herbert Simon, “Theories of Decision-Making in Economics and Behavioral Science”, *American Economic Review*,44:272）为什么不在实验中增加合作模式的因素，产生随机性和限制性信号，从而让实验客体有机会建立合作认知模式，而非仅仅的随机性呢。假设为了更好发掘实验客体的潜力，我们可以增加一个实验客体，他的收益取决于其能够截取并修改信息的合作伙伴，这样我们就可以得到与穆尔相似的结果。增加了实验中的因素后，我们可以更好地验证客体的行为模式建构，这有助于解决相关的心理学、美学领域的一些复杂问题。西蒙在上面的文章中指出（第 426 页），计算机程序可以“通过图像或模拟”证明几何原理。这具有非常重要的现实意义。（这促使我们相信零和博弈的“人性恶”这一假想并不适应于数学创新。自然的各种现象向人类揭开其神秘的面纱，避免人类艰难地探索。）

则这一推理逻辑可能逐渐与客观现实相脱离。例如，一个人投票选举某个候选人，并不是因为他想主动接近大家的立场，而是当大家都这么做时，他别无选择。大家心里都知道——这不是预测瑞恩高德（Rheingold）小姐当选，而是如同投资存在增长潜力的绩优股或房地产——每个人都希望别人也期望别人去投资。投资珠宝也是同样的情况。不过，最典型的例子是黄金的货币作用，这一现象只能被看做是协作博弈的解决方案。（最常见的家庭版协作博弈是当电话突然断掉时，如果双方随后同时打给对方，就会占线。）

现在，我们回到选择“正数”的游戏。第3章中的实验表明，大多数人都会选择3、7、13、100 或1这样的数字。但是，当要求他们选择与对方都感兴趣的相同数字，或当他们也知道别人在做同样的事情时，每个人的出发点都会不同。其中，最有可能的答案是1。最好的逻辑是：不存在大家都喜欢的数字。大家的选择多种多样，可能是3，也可能是7，或者是别的数字，没有好办法选择一个最受欢迎或最醒目的数字。这时，如果每个人都问，这些数字中哪个最醒目独特或者怎么才能选择到大家都喜欢的正数，他也许会猛然醒悟，最有可能的数字就是“第一个”或“最小”的那个正数。[1]

协作问题的博弈默式图4-1表示完全协作问题的结果。一方只能选择一行，另一方只能选择一列。

双方的收益取决于他们选择的数字。如果一方的选择与另一方的惟

[1] 这里我们可以引用英国经济学家凯恩斯（Keynes）经常引用的案例。这个案例证明，它不仅能够解决书中出现的问题，对“解决方案”的解释往往并不惟一：专业投资如同报纸竞争。在后者，竞争者们必须从数百张丽人照片中选择六张。谁选择的照片能够赢得其他竞争者的普遍赞许，谁就获胜。这样，每个竞争者都会以其他参赛者的审美标准，而非自己的审美标准选择照片。结果选出的照片不是选择者自己最喜欢的，却是其他参赛者最喜欢的。现在，我们达到第三级别，在这个级别上我们得到了普通人对其他普通人的观点。我相信，一定会有人达到第四、第五甚至是更高的等级（Keynes，*The General Theories of Employment,Interest and Money*，New York，1936）。这类博弈游戏表明，参数行为与大数字的一般关系不适用于存在众多平衡参数的默式博弈。如果要求其他行为具有参变性则要求这些行为必须具有可观性，而非猜测。不管存在多少对手，默式博弈的非参变特性保持不变。

1	0	0	0	0
0	1	0	0	0
0	0	1	0	0
0	0	0	1	0
0	0	0	0	1

图 4–1

一选择相同，二者都将得到奖赏。我们将所有的收益分布于一条斜线上，并且将数字 1 定位于这条斜线上的空格内，其余空格为零。（每个空格代表一个数字完全不影响我们对双方收益的实验分析。）

实验过程中，我们必须排除其他相关理论原则的干扰，即［以卢斯(Luce）与雷弗（Raiffa）的话说］，对行列进行标注或选手双方得到相同的收益。[1] 因为从某种程度上，这样做的结果是战略被人为地做出标识——即能够超越博弈数学结构的一些符号或暗示——从而导致双方选手可能充分利用偶然机会"赢得"游戏。由于同样的原因，这些博弈游戏本身具有趣味性和重要意义。

甚至在图 8–1 中，博弈游戏的行列也可能具有一丝的暗示作用。如果正如收益矩阵所示，尽管双方选手努力避免过分依赖偶然因素，这个矩阵博弈游戏似乎也非常容易"取胜"。（如果将矩阵分为无限个行和列，那么双方选手将更容易而非更困难取胜。如果这样，这个游戏更接近于前面的选择"正数"游戏，只是二者所作的标识不同而已。正因如此，才不会有少数人选择 3、7、13 等数字。）鉴于矩阵更多地关注"第

[1] 在卢斯和雷弗的案例中以及纳什（Nash）的对称性假设中都排除了对方标符的使用。参阅 Nash, "The Bargaining Problem," *Econometrica*,18:155~162 [1950] and "Two Person Cooperative Games" ,*Econometrica*,21:128~140 [1953]。一般形式的博弈实验都排除了对默式或显式非零和博弈的战略标符因素。上面的收益矩阵出现的是他们的抽象化。（收益矩阵仅仅是一种分析方法，而非博弈实验的一部分，因此不存在实际战略的左右、上下等排序问题。）前面提到的打电话问题突出表现了该问题，即谁打回电话，谁等候电话。

一个”、“中间”、“最后一个”等数字[1]，所以仅仅建立这样一个矩阵对双方选择并非完全公平。如果各种策略方法无法以数字或字母等排序，而分别被随机冠以一个特别的名字，那么，这些名字一定能够协调双方的选择。

这样，我们逐渐清楚，完全冲突策略或协作战略的选择方式多种多样并不惟一。至少当一方承认零和博弈中随机性的“最大最小”解决方案时，情况是这样。在完全合作博弈中，双方的目标是希望通过某种直觉或双方间共存的某种暗示实现沟通。然而，在完全冲突博弈中——随机选择的典型——双方的惟一目标就是避免与对方存在共识，甚至是偶然形成的某种共识。[2]

为了进一步说明这个问题，我下面再举一个例子。假设我从一副

[1] 这一观点是作者早期发表的实验结果之一，在默式博弈中，有关“依赖于独立选择”的假设毫无意义；同样，这一假设也不能适应于显式谈判博弈。潜在的结果取决于双方能否做出共同的选择，尽管情况有可能并非如此。有关探讨请参阅本书 127 页的卢斯和雷夫的案例。

[2] 随机化战略也可能无法实现选票在候选人之间的合理分配。假设 100 名选民中，55%的选民出席，并知道这一事实；或者假设 6 名候选人中的两人选票非常接近；或者三名得票最多的候选人将成为董事会的董事，那么就可能出现大部分选票将集中到第一个（或第二个）最集中的选择，并使另外两个得票较少的获胜候选人分别得到 22票。但是，如果多数派中的每个人都以掷币决定将选票投向自己党派的候选人，其中候选人只有 1/6 的概率得到至少 22 票。如果少数派没有明确方式实现合作或者依靠一个随机机制，那么多数派取胜的概率将非常大。

一个部分随机化战略也可以被用来减少冲突领域。假设分别坐在牌桌的北侧和东侧的两个人打算坐到附近的另一张桌子上，而且必须在缺乏沟通的条件下选择位置。在这种情况下，如果他们成功地找到了自己的位置，二人分别得到 1美元奖励。这是一个简单的协作问题。尽管如此，我们为了扰乱双方选择动机，决定当双方成功地坐到一起时，位于右边的女士们将额外得到 2 美元奖励。现在，这个博弈场景失去了平衡点，双方缺乏利益共同点。在这种情况下，任何座位安排都无法使二者产生换位的意愿。（尽管双方都期望自己坐在对方的左侧，但是却无法实现。）一个随机战略使双方各自得到 1美元。但是，如果一方决定在利益完全共同的博弈场景中选择自己的位置，并选择掷币决定自己的位置，并保证双方都不能不同选择同一个位置或者坐在自己的对面，那么双方将得到同样的奖励。这是一对平衡的（混合）战略选择，双方应该分别得到 2 美元奖励。

52张的扑克牌中抽出一张，现在让你猜测我选了哪一张。传统博弈论关心的是，我如何抽牌你才不能猜测到。我可以任意抽出一张，那样你除了盲目地瞎猜外别无选择。而现在我们讨论的博弈论关心的是，我非常希望你能猜出我抽的是哪张牌。而且，你也知道我会配合你猜出是哪张牌，那么，这种随机抽牌的方式显然无法帮助我们达到目的，即无法实现我们之间默契合作。福尔摩斯也许会通过掷币来决定在哪个车站下车以削弱特定车站的标识作用，因为大多数情况下对方只有 50%的猜对机会。但是在利益共存的情况下，他们需要使用车站的标识而非纯粹的侥幸。而如何做到这一点则更多地依靠想像力而非逻辑推理，更多地依赖诗情画意或幽默感而非数学推理。需要值得注意的是，传统博弈论并没有给予这一博弈场景应有的“评价”。传统博弈论认为，示例分析方法——或许系统分析能够做到——根本无法回答人们如何以这种方式达到共识。换句话说，这种博弈观点从根本上依赖于经验证据。[1]

读者需要注意的是，强调“标识”（博弈游戏中具有象征性和启发性的细节）的作用和该理论对经验实证的依赖并不涉及博弈论是否是先证性的标准化理论问题——涉及实际决策的广义性（generalization）或

[1] 在这种情况下，我们需要考虑的不仅仅是两方一定信息共享的代价，还要考虑什么信息模式产生合作，以及在马塞克（Marschak）的小组理论（theory of teams）中为自己找到证据。

这里有一个类似博弈论中的“囚徒困境”的情景。两个同伙在实施犯罪之前被逮捕，并被警方分开审讯。双方必须口实一致，否则都将被判刑。但是警方给他们提出一个条件，坦白从宽，抗拒从严。这样每个歹徒都面临最大最小选择，不仅要考虑符号这一条件的犯罪事实，而且还要考虑同伙的口供是否与自己的一致。这样他们之间展开了一场博弈游戏。具体结果如下：

0.5 0	0.5 0	0.5 0	0.5 0
0 0.5	1	0	0
0 0.5	0	1	0
0 0.5	0	0	

（左下方是选择行的选手的结果，

右上方是选择列的选手的结果）

正确决策的策略问题。之所以如此强调并不是说，人们简单地受到符号标识的影响，而是他们为了正确参加游戏必须受到符号标识的影响。更重要的是，一个标准化理论必须有利于策略的制定或至少产生其他积极效果，即使人们没有借用这一理论。标准化理论不能否认或排除对博弈双方有利或双方共同利益必需的游戏细节。争夺舞池的两对夫妇或争夺停火线的两支军队，在决策过程中都会受到现实条件的限制，从而给双方造成不必要的伤害或损失。

只要我们承认理性博弈选手达成预期共识的逻辑方式，“标准”(normal，数学抽象形式）的博弈游戏与“广义” (extensive，具体形式）的博弈游戏在逻辑上存在差异这一事实完全可以证明这一普遍规律。正如第 3 章中所指出的，显式谈判也存在相同的条件因素。描述这些条件因素的传统术语“非合作性”根本无法准确表达默式博弈的含义。它以自身独特的方式发挥合作的作用，而且即使在我们增加冲突和默式混合博弈等因素后仍能发挥其原有的作用。（附录 C 中，有人认为博弈中的一些相关解决方法可以按照协作思想被重新释义。)

混合博弈中的暗示与互动认知

合作博弈理论之所以能够引起人们的关注，不仅在于其自身独特的魅力，更在于合作博弈理论揭示了混合博弈的本质。尽管在完全默式博弈中，博弈双方无法进行有效的沟通和交流，但是完全默式博弈中还是存在许多合作因素。为了便于大家理解，我们这里向大家介绍一个与第 62 页的第六个问题基本相同的博弈游戏。

一个选手“位于”辛辛那提，另外一个选手位于旧金山。两人手中都有一张相同的美国地图，并打算划分美国各州。每个选手都按照自己的意图划分，分界线可是一条笔直的线，也可以是一条曲线，划分过程中不必考虑地理或行政因素。如果划分的结果不同，两人都将一无所获；如果两人划分的结果相同，二者将得到各自范围内的财富，即各自所拥有的城市。现在我们还不能将奖品具体化，他们的标准可能是领土，或人口，或工业财富与农业资源等等，二人也有可能得到不同的奖品。换句话说，即使整个领土都具有巨大的价值，也并不意味国家的每

一部分都有价值。总之，根本不存在统一的价值标准。（结果，二者之间根本不存在实现价值对称的划分方式。）

在这个博弈游戏中，一个最关键的问题是双方必须实现合作。一方只有按照对方的期望行为，并知道对方也将按照自己的期望行为，他自己才有可能取胜。他们必须一起找到一个实现双方利益共享的分界线。任何一方都不能轻易地“故作聪明”，自欺欺人地欺骗对方。

第3章中的实验表明，选手双方面临这种博弈游戏时并不是孤立无援的。游戏也并非像无限划线那样“无限”困难，因此游戏的各种参数变量并非难以控制。但是，一个成功的结果取决于完全合作博弈的众多关键因素。实际上，双方选手在收益相同而非利益冲突的条件下，将主动协调二者的立场和条件，从而在类似的博弈游戏中取胜。关键在于他们能否在游戏中一起找到某种符号或暗示或理性化条件，并且自愿接受这些因素的约束和制约，即使这些条件对其中的一方并不是十分有利。他们必须努力找到这些暗示条件。（例如，虽然地图有很多地方可以作为双方的暗示标识，但对双方而言，很难找到一个满意的暗示标识。这时，他们需要一个中间人或调解人划出一条界线。无论公正与否，他们都必须接受调节人的意见。）

但是，在某种条件下，特别是在没有冲突的条件下，协作因素与沟通问题存在密切的联系。如果博弈双方一帆风顺、坦坦荡荡地达成共识并实现合作，那么，完全合作博弈不仅增加了其自身魅力，也提高了其作为博弈理论的价值。现在我们不禁要问，在众多显式谈判存在的环境中，协作因素在混合博弈中究竟占有多么重要的地位？

对协作原则的深入研究来源于两个重要方面的考虑。一个是在第3章中讨论过的。默式谈判为我们提供了一个分析问题的方法——或许是一种类比，或许是实际心理思维活动的一种反映——在完全谈判过程中，通过一个“理性”过程寻求博弈双方的协作。因为双方都认识到，即使达成一个略具偏颇的结果总比没有任何结果要好得多。默式谈判所具有的并发挥巨大作用的心理现象“互动认知”，在分析显式谈判时也具有重要的作用；预期协同（*Coordination of expectations*）就发挥着同样的作用。

其次，我们所要分析的大部分谈判过程或博弈场景，都或多或少地存在默式谈判的因素。在一些场景中，如前面提到的交通堵车案例，双方根本无法进行有效的沟通；而在另外一些场景中，如发展邻里间友好关系的案例中，双方有时出于保护隐私的需要，根本不愿意进行沟通。在非法谈判或外交谈判中，窃听对方的信息对双方而言都是一件令人尴尬的事情，更不用说你来我往、飞鸽传书了。假设博弈游戏涉及多方选手,如以种族界线（ racial border）划分住宅区和工作区所进行的谈判，也许就没有现成的机制规定适合于显式谈判。这种情况下，虽然沟通是谈判重要的组成部分，但实际行为也不可或缺。因为与其说是一场对话游戏，还不如说是一场“ 行为 ”游戏。

而且，如果双方之间存在“ 活动 ”的空间，那么双方就会抓住这个有利条件在谈判期间进行必要的活动，特别是一些互动行为的具体作用需要经过一段时间方能见效。当然，这不是要求谈判双方一开始就一定要进行相关的活动和互动行为。如果那样的话，再加之双方之间必要的对话和沟通，谈判的进程无疑会加快。如果这些互动行为仅仅具有某种暗示作用或提示作用，我们就可以将其划入沟通过程中的对话范畴。但是通常情况下，即使类似具有暗示作用的互动行为也发挥着重要的作用，会对谈判进程产生重大影响，甚至他们的暗示作用远远胜过沟通过程的纯对话。有人会说，不见棺材不落泪，你不打枪怎么知道枪的威力；也许有人说，他发现了一块具有战略意义的地皮，但是只有投资、冒一定的风险保护这块地才能证明其战略意义。互动行为可以揭示双方的价值观或其选择的空间；当对话无法发挥应有作用时，互动行为却可以促使谈判双方采取必要的行为。此外，双方不必拘泥于在谈判会议上的某种形式，单方面的努力也可促进互动进程。

换句话说，双方之间的谈判博弈是一个动态的互谅过程，而并非某个协议僵化规定的完全沟通过程。在有限战争中，战争双方斗智斗勇寻求战争有限条件的案例充分说明了这一点。为了进一步说明这个问题，我们下面将向大家介绍一个案例。

默式博弈游戏[1]　假设我们的两个选手各有一张相同的美国地图，在此之前，每人得到 100 个硬币，并被告知游戏规则。现在每走一步，双方都要从自己的 100个硬币中拿出 5个放在地图的州标识上面。如果双方各将 1个硬币放在同一个州，那么这 2 个硬币都将被裁判拿走。如果一方将 1 个硬币，另一方将 3 个放在同一个州，那么双方都将被拿走放在该州的 1 个硬币，留下 3个中的 2 个。下一步骤同样。如果 A 将 2 个硬币放在 B 放过 1 个硬币的州，那么 A、B 的将被拿走 1 个硬币，留下 A 的另一个硬币作为其占领该州的标志。依此类推，直到双方用光持有的硬币为止。接下来，一方可以从自己占领的州上凑足 5 个硬币分布到其他州。如果二者放在同一个州，重复上面的原则。这样，直到双方通知裁判二者愿意结束游戏。

现在，游戏双方将根据比赛的结果得到自己的奖品。双方每在地图上留下 1 个硬币，将得到 1 美元。此外，还要参考其占领的州数，包括放有硬币的州及其包围的空州。

之所以在奖励选手时考虑州数，主要是为了强调 48 个州的货币价值，也可以说是它们的“经济价值”或其他类似的说法。没有人能预测到双方得到的最终价值相同，甚至是非常接近。对于一方来说，也许人口是州价值的重要组成部分；而对另一位选手来说，人口可能就并非如此重要。双方都不知道对方的价值观或者说略知一点而已，例如一样东西在对方看来具有多大价值。双方只能通过互动行为认识各自的价值观。

这里我们向大家介绍的是，一个以互谅为基础的混合博弈，游戏的进程取决于双方的互谅程度。如果双方无法达成互谅或者互谅意识薄弱，那么双方将同时受到一定的损失；如果双方都无法预测对方本次将

[1] 第 6 章将证明这样的案例不仅可以作为具有说服力的论据，而且具有重要的研究价值。在非零和博弈游戏一开始涉及如何激励选手的问题时就应该引起注意。在零和博弈中，胜利涉及一方近期的对方，智慧挑战和双方的促使选手找到正确（惟一）胜利方式的竞争激励机制。但是在非零和博弈中，胜利并不意味着一方的完全胜利，胜利的目标也不是将对方完全打败。除非双方得到真实的奖励，否则在游戏中通常涉及第三方，游戏的结果也是相对优势而非绝对优势的一方（这也是为什么没有二人的非零和室内游戏的原因）。

把硬币放在哪个州，那么他们就不会为了拔城夺寨而蒙受损失；如果一方想夺取对方的州，就必须为之付出至少一个硬币；如果对方决定重新夺回推动的州，他就必须为先前的行为付出更大的代价。双方不仅要失去被裁判拿走的硬币，而且还要为被占领的州占用一部分硬币。当游戏结束时，他们如果没有足够的硬币进行游戏，那么地图上一定会留下为数不少的未占州。

现在，双方如何在游戏中进行“谈判”呢？可能是一种方式，也可能是其他方式，实际上，双方都会不断提出条件与反条件。他们或接受对方的条件，或拒绝，或报复，甚至寻求威胁或允诺。[1]假设我们剥夺他们进行言语沟通的权利，那么他们为了传递信息只能付诸于某种互动行为。一方必须注意观察对方的行为举止，而且必须想尽办法准确地向对方传递自己所要表达的信息。如果一方发现一个谈判的某个阶段对其具有十分重大的价值，他可能会在此处徘徊不前，与对方纠缠不放，甚至不惜损失几枚硬币与对方两败俱伤，直到对方放弃为止。因此，双方最好在发生误解之前能准确地了解对方的迫切需要和利益战略点。如果一方决定妥协让步，以国家的一大块土地换取自己迫切需要的部分，那么，他不仅会主动放弃一部分城池拱手让给对方，而且会以自己的疆土格局划出一条界线。

但是，这条疆土格局是如何形成的呢？他们双方并非十分了解这个游戏的数字结构，况且我们人为地制造了许多障碍，如双方都对彼此的价值观缺乏了解，根本无法实现所谓的对称性或公平性或巨大帮助。假设他们在游戏之前知道自己在某些领域的疆土格局，如他们价值观中已经形成的自然边界、熟悉的政治环境、每个州的经济特点、以及在游戏过程中建构的完形心理学（Gestalt Psycology）、陈俗陋规。[2]

显式沟通　现在我们改变一下规则，允许选手畅所欲言，随意进行

[1] 前面的游戏中也涉及了相关问题。

[2] 如果邻居的果树伸到了我的后院，我把墙这边的果实全部摘了，那么邻居就会知道我可能提出的“条件”。如果对方默认的话，那么双方就会继续接受这一分配方式。相反，如果我摘了果实总数的一半或按照自己家人口摘了，那么邻居可能无法知道我是怎么想的。（或许假设我只随意摘了墙这边的一部分，那么邻居可能会更倾向于报复或抵制行为。）

对话沟通。那么这将对游戏产生什么影响呢？这在某种程度上有利于双方进行沟通，提高游戏的效率；促使他们达成某些共识，特别是在默式条件下无法实现的交易；或许他们可以避免就一些州进行争执，从而避免两败俱伤。当然，我们无法保证双方没有你争我抢，夺城掠地，碰得头破血流的情况。特别是在双方能够进行对话的有利条件下，首先发盘的一方更不会放弃这一绝佳良机。除了双方各自的游戏方式，没有其他方式能够说服对方。（我们现在让双方告诉对方自己对每个州的评价，而且允许他们说一些无伤大雅的谎言，但是我们同时也不会为双方留下任何有关他们个人价值观的一纸凭证。）

由此可见，对话沟通因素的引入并没有给这个游戏的属性产生巨大影响，只是一定程度上影响了游戏的某个具体结果。和此前一样，他们仍然需要彼此间就双方观点立场进行交流和沟通，同时需要某种程度上的互动行为以及就相关规则和条件的默契配合。

这个游戏的最显著之处在于其与零博弈和零和博弈涉及的最大最小解决模式形成了强大反差。根据最大最小方式解决问题，最终将导致零和博弈归咎于一方行为。一方不仅没有必要与对方进行沟通，甚至不需要知道对方是谁或是否存在。随机化战略更对沟通恨之入骨，处心积虑地排除双方进行沟通的一切可能性。随机化战略除了博弈结果的数学结构外，也不允许一切博弈细节存在，更不用说博弈双方之间的沟通因素。

在国际象棋中，棋手双方不在乎局中的棋子是否看起来像马、象或者其他东西；或者这个游戏本身是否叫做“象棋”、“内战”还是“实业”等；或者棋格被扭曲像个行政区或地理区域。棋手双方是否认识对方、讲同样的语言，或拥有同样的文化背景都不重要，更不用说棋手的名字以及谁给起的名字。（如果这些因素都重要的话，一方完全可能人为地破坏这些细节的影响。如果需要的话，任意一种随机的最大最小战略也会消除这些影响。）

但是，如果我们改变象棋的收益矩阵，使之成为一个非零和博弈，最终的胜利不仅取决于双方消灭对方的棋子数，还要参考双方留下的棋子数以及他占领的棋格。通过这种方式，棋手双方都不会轻易地吃掉对方的棋子，因为这也可能给自己造成损失。同样，这种方式也使棋手双

方都不知道对方的价值点，即哪个棋子和哪个棋格对对方具有重要价值。由于时间的限制，双方都不可能故意与对方搭讪以拖延时间。

而现在，假设以下任何一个因素的出现都将对双方产生重大影响：我们是否称其为“战争”和“黄金潮”；棋子看起来是否像马、战士、探险者，还是像寻找复活节彩蛋的孩子；棋盘的背景是地图还是图画以及棋格被扭曲成什么形状；或者在棋局开盘前介绍双方的背景故事。

这里我们再次修改游戏规则，要求双方必须以他们自己的方式达成共识，无论采取哪种沟通方式（言语和／或行为）；无论以何种方式（默式和／或显式），双方必须自定游戏规则，进行有效沟通，并努力达成某种共识和避免双方利益的共损。其中，一些“偶然因素”也许有利于双方尽快找到彼此的行为模式或游戏中的一些隐含条件，而且双方制定的规则必须有助于区分各自的收益。双方必须这么做，因为这有助于而非阻碍双方选手发掘彼此稳定清晰的互动行为模式。虽然最根本的思维过程和心理反应都是对传统的创新，但是传统创新的必要因素和发掘传统潜在优势的素材与游戏的数理内容并非总是相符。[1]

游戏结果取决于双方对彼此行为的预期判断，而且彼此也知道双方的预期存在互动性。选手双方必须通力合作，形成默契，找到双方的共识和决定游戏结局的比赛方式。他们必须制定共同的游戏规则，并接受由此产生的游戏结果。

最典型的沟通问题是，当博弈一方突破另一方认为的“界线”时，后者如何克服意欲对前者实施的报复行为。假设双方之间缺乏有效的沟通，那么一方的报复能力不仅取决于形成“界线”的语境，还取决于另

[1] 一个典型问题就是如何划清原子武器与其他类型武器的界线。根据现在的研究成果，如果仅仅以威力作为判断的标准，那么很难将二者区分开。但是，如果人们认为存在一个界线的话，他们也会这么做。这个结果完全是人们的一种假想：十年来，人们一直认为二者之间存在界线，而且大家都这么认为，甚至持不同意见的人也只能将头紧紧地缩回去，也许只有当另一场大战爆发时，人们才会醒悟。这完全是一种传统的思维方式，如同将罪犯关进监狱并非是“残暴”的惩罚或十多年来一直将议会里的大学代表看做是英国民主的表现。和大多数传统一样，随着时间的流逝，人们对于原子武器的这一看法也许会改变。（在附录 A 中将就相关问题进行详细论述。）

一方对这一报复模式的认知能力。如果他见到过类似先例的话。历史和文学先例、司法和道德案例、数学和美学以及熟悉的生活常识都影响到一方的报复模式或对别人意图的认知。即使双方能够进行有效沟通，这一状况也不会有很大的变化。事实胜于雄辩。

游戏中的暗示性细节对结果的影响、选手对暗示条件的依赖以及游戏的提示性标识，都不仅仅涉及对非零和博弈选手行为的研究。虽然问题的焦点不是选手应该根据游戏的数理特征做出反应，但是他们必须考虑这些因素。因此，甚至是标准性理论——博弈战略论——也必须承诺，理性选手可以共享这些有利条件；即使认识到这些条件分布相对的不合理性，双方也必须理性地接受这一现实——对方也会理性地期望另一方接受游戏中各种细节的暗示。为了避免两败俱伤，后者也会这么做，给予对方充分的合作。[1]

模拟实验　为了深入说明作者的观点，下面我们引入一个模拟实验。（希望类似的实验能够付诸实践。）该实验类似于概念类比，对谈判双方的心理现象进行实验检验。

在实验的第一阶段，我们必须发明一种类似于测谎仪的仪器用于记录和测量实验客体的“认知”程度，也可称之为关注程度、警觉程度或

[1] 此外，应该强调的是心理学家进行的相关实验，证明了在谈判过程中或完全合作博弈中的关键点问题和内禀磁矩问题。例如，即使有些人存在一定视力缺陷（并非完全失明），他看到的仍然是一个完全的世界。但是，其中有些图像是他们简单地“完成”的；而有些图像却是因为熟悉而“完成”的。而图像的产生却并非这么简单。科夫卡（Koffka）指出：“我们的周围并非都是长方体，但是我们看到的却都是长方体，因为纯长方体比一个模糊的长方体要好。”观察静止过程的最大最小属性后，科夫卡认为心理过程也应该具有这些属性：“因为至少我们可以选择简单条件下的心理组织进行研究，从而断送它们一定具有规律性、对称性和简单性。”这一结论是建立在类质同像原则基础之上的。根据这一原则，心理过程的特征也是认知过程的反映。这样，我们可以得到一个普遍的规律指导我们对心理组织的研究，尽管有些模糊。这一原则可以被定义为：“心理组织和现在的条件一样‘有利’。这里原‘有利’并没有特指，或许是我们在争论中常说的规律性、对称性或简单性。”

如果个人认知与各种形式的组织都受到这一条件的限制，那么，涉及预期协同的互动认知或互动形式组织一定受到类似条件的限制。鉴于非零和博弈要求最终的共同形式组织，因此可以说标准战略论（不仅仅是描述心理学）一定也受到相关条件的限制。

兴奋程度。我们还需要一台仪器用于测量，在实验客体浏览一组有序的可能性结果过程中，什么样的结果能够引起他的注意以及其在实际谈判过程中表现出来的兴奋程度。

现在，我们具备这样一台仪器，并开始我们的谈判实验。为了方便操作，我们允许双方通过达成某种共识实现部分收益。同时，我们为博弈双方提供足够的“讨论话题”，为双方提供充分的争论、辩论和选择的空间。总而言之，我们的游戏不仅仅是一个具有明显中分点的数学模型。

现在，我们将实验双方与测量仪器连接起来，并保证双方能够看到彼此的仪表数据。换句话说，双方都能够知道彼此对某个结果的反应。我们开始运转这台仪器，将一组可能的结果数据以有序或无序方式不断地向双方展示，每个结果停顿片刻，并亮起提示信号。同时，我们保证双方在浏览数据的同时，能够看到彼此的仪表，甚至可以看到对方的表情，如果他们愿意的话。

最后，我们完成了整个实验过程。但是，实验过程存在众多偶然性。实验有可能排除了显式谈判，而仅仅是数据的反复浏览。我们需要观察双方的心理反应数据最终是否会集中到某个结果，即双方在浏览数据的过程中，对某个特定结果所表现出来的非自愿性相同反应。（为了实验的需要，我们在实验过程中也许会让一个选手回避，另一位选手单独浏览数据，从而得到双方不同的反应。）如果双方能够产生某种共识，我们将得到两个非常重要的相同现象，无论我们能否称其为双方的心理谈判进程。我们得到的结论是：（A）双方确实受到谈判条件的影响。(B) 双方的反应取决于二者之间的互动，即一方能够通过仪器显示的数据看到对方的反应，并知道对方能够通过仪器数据看到自己的心理反应。［作者认为，和洛特（Lot）的妻子一样，双方选手往往也无法避免自己对某个结果的关注，甚至是对自己不利的结果。[1]有意识地回避

[1] 尽管科夫卡引用的例子令人难以置信，但却切中要害。“一个专家注意到，足球守门员往往能够拦截住对方偶然的进球，尽管对方考虑到守门员可能会将球拦住。守门员往往在空中形成了一个转移对方进攻球员的注意力的关键点。如果对方快速运球的时候仅仅盯住守门员，那么他的球很可能落在守门员的附近。但是当进攻球员重新组织进攻，改变守门员的重心而转移到另外一个点，那么他还会重犯刚才的毛病。”

“关键点”通常反而更能增强关键点的“关键”作用。]

另一种情况是，我们可以让双方选手浏览数据的同时进行公开谈判，并保证仪器在谈判过程中继续对双方的心理反应进行跟踪记录。（我们允许双方选手，如果他们愿意将其作为一种战术的话，可以引用仪器对双方心理反应的记录数据。例如，一方完全以可以以对方的血压数据为依据，对另一位选手说：“你的条件完全不是60美元，因为你的血压数据告诉我，你能接受的条件是40美元。）

这一实验可以得到三个结论。首先，我们可以通过某种方式，知道一个选手在决策过程中表现出来的对选择条件的自然反应，并且不同的选择条件能够产生不同的效果。其次，如果一方知道自己的反应对对方暴露无遗时，这些反应数据都可作为谈判的暗示条件。某种程度上，双方之间清晰可见的心理互动可以被看做是双方在实际“谈判过程”中的互动。最后，实验得到的可测性现象，也可以说是我们设计的谈判过程，一定程度上是或涉及我们通常所指的谈判过程，至少与其存在一定的联系。（类似的实验如果涉及多人博弈，将会更奇妙。）

这个实验并没有经过实践的检验，因此不能作为引证的论据。可以说，这个实验是作者意欲建构的理论系统的模型。作者希望借此证明自己理论体系的合理性：谈判不仅涉及双方的预期协同，而且谈判过程最终形成的这种协同可能依赖于谈判过程自身的动态因素，而非仅仅是谈判的初始条件。

焦点式方案的动态特征　焦点式方案本身具有的特点，不仅将其在本质上区别于其他相关选择方案，而且呈现出许多动态特征。例如，与其他方案相比，焦点式方案更不容易让博弈双方做出较小的让步。从另一个角度看，这意味着该方案产生的结果十分明确，毫不含糊，从而更具有说服力。假设谈判一方坚持的条件是50%，即使失败了，他也不会接受对方47%的条件，因为一个微小的让步意味着彻底的失败。此外，双方在原则性问题上通常很难妥协，而焦点式方案往往关注双方的原则性问题。我们不能指望以领土上的让步来满足侵略者的侵略欲望，因为对方知道双方都期望我们让步，直到找到一条“合理”的、具有说服力的新国界。

事实上，协议的关键点往往取决于其协议自身的特点，即不允许妥

协的存在，即使是非常无关紧要的让步，因为一小步的让步可能导致一大步的损失。一方也许会在某个边界划一条线或将自己的立场建立在不断重复的原则上，“如果不是条件，还会有其他条件！”一方立场越强烈，妥协的可能性就越小，焦点的说服力就越大。这与我们在戒烟或戒酒过程中经常与自己进行斗争具有同样的道理，当我们想喝酒或吸烟时，总会对自己说“仅仅一点而已”。这无疑是一个小小的让步。尽管如此，这种情况毕竟是少数，大多数人还是宁愿戒掉烟或酒而不想做出让步，因为一旦有了第一次，就会有第二次、第三次甚至更多，结果戒烟或戒酒的努力功亏一篑。所以，一旦人们认识到了这一点，戒烟或戒酒就会相对容易些。

有时，焦点本身也并不稳定。如果那样的话，它所发挥的作用不是代表一个结果，而是提示双方结果的位置。议会中的试投票或者选手双方在继续进行的游戏中的互相探底都是同样的道理。就本质而言，焦点的确定对双方也是一个挑战或大胆的举动或拒绝行为，即要么迫使对方做出必要的让步或者自己做出必要的让步。这只是代表博弈游戏本身的博弈游戏的一部分，即帮助双方建立预期模式。有时，博弈双方将其作为一种战术人为地加以利用；而在其他情况下，这只是提示双方避免妥协的客观符号而已。

这样的例子还有很多，诸如对红色中国的外交承认、美国大学里的宣誓行为、重要行业罢工的解决、将鸡尾酒酒会的场地让给一个不速之客以及政治大会上就某个问题的投票等。有时，一个问题的解决能够为其他问题的解决提供了思路，如议会里就某个问题进行的民意投票能够说明该问题的影响。但是情况往往是，一个问题并不能代表所有的问题，其他的问题可能需要的是某种暗示，如默认。只有这样才能证明，哪一方从双方的预期互动受损，或哪一方受益。

在协作博弈中，这一现象往往被看做是实际暗示符号。在一个缺乏有效组织的同盟体系中，其成员往往能看到同盟成员共同行动的可能性，但不能肯定同盟成员一定能够采取共同行动。每个成员都想知道别人将要采取什么行为，以及他是否采取自己所预期的行为。议会中的民意投票（test vote）或组织的某个集体行动，如大规模抗议，都证明某

种形式的同盟的存在，并表明大家都期望共同行动。但是，甚至在二人博弈中，双方心理优势或心理让步最终都将反映于最终结果中。

实验心理学研究表明，双方在博弈游戏中进行的各种互动行为具有符号意义，并能够为双方实现预期协同创造条件。这一发现无疑是对博弈论的重大贡献。

数理方法的实际应用　我们必须注意到，研究分析人员得到的结论与博弈选手的实际结果或许并不完全相同；或者说，对实验人员具有重要价值的线索，对选手而言可能却毫无价值。特别是复杂的数学方案具有的博弈特点（除非一种简单的其他方式也能得到同样的结果）不一定能产生焦点式预期判断或影响博弈的结果。纳什等人的结论只是来自于他们的经验释义。与美学属性、历史属性、司法和道德属性、文化属性以及其他属性一样，博弈的数学属性有助于某些博弈选手预期判断某些结果。如果选手双方本身就是数学博弈的理论家，他们可能都会认识到这一点，并受到具有数学属性的潜在解决方案的影响。双方都会努力排除偶然因素的干扰，并明白对方也会这么做；而其他非数学家的博弈选手往往过分地关注这些影响双方预期判断和结果的偶然因素。

（在大多数情况下，数学通常具有惟一性和对称性；而后者则不一定具有数学的精确度和要求；或者与双方通常以理性的非数学方式得到的、具有本质属性的结果相吻合。）

尽管数学方法只是能够影响双方预期判断的博弈方法之一，但是它们和其他方法一样也涉及同样的心理机制——影响双方实现预期协同。前面出现了相同道理的案例，如在百货商场失散的夫妇能够默契、自觉地到服务台等候对方，因为双方都认为这是彼此心目中“醒目”的地点；同样的道理，博弈双方都是数学家选手——双方都知道对方是数学家——就会努力以数学方法解决问题，而非依靠言语的沟通。

这并不是说，双方在博弈论的规则下一定要具备足够的数学知识。我们这里关注的是，选手双方相互之间的理解、互谅以及双方在游戏过程中对各种提示性条件的释义，而不是双方在必要时达成和局的条件。假设双方谈判达成协议的过程仅仅是一个心理沟通的历程——实现预期判断的共鸣——那么，我们可以认为数学博弈论并非实现共识的必要条件。因此，

我们可以做出这样的结论：至今为止，尚未有任何充分证据表明，数学方法是实现共识的主要方式（附录B将对这一问题进行深入探讨）。

或许有人不同意对谈判者预期判断形成方式的假设，不论是在谈判过程中或是在谈判之前，还是受到谈判本身或其他因素的影响。但是，我们必须看到，一些预期协同现象直接、迅速、准确地预见了谈判的结果。不管是以公开方式表示同意（或反对）还是默认某个结论，他都必须（如果他有自知之明的话）承认自己做不到这一点以及别人也将报以同样的想法。博弈双方通过协作达到共同的目标，这一事实足以验证预期协同的作用。

心理沟通　混合博弈对双方在互谅博弈中的行为互动的重视进一步提高了其自身地位，并与零和博弈形成了鲜明对比。在零和博弈中，博弈双方之间是完全的零和关系，一方不可能准确地知道另一方的价值观；而在混合博弈中，互动行为意味着双方间的心理沟通。

迄今为止，我们还无法建构一个放之四海而皆准的谈判游戏。在这个谈判游戏中，双方能够完全了解对方的心理活动。假设双方知道对方的选择矩阵，如同猜测博弈游戏的机制结构。原因在于除了在特定的条件下，一些谈判主体根本不知道谈判过程中的某些因素。我们怎么才能知道，苏联人如此厌恶的一场两败俱伤的毁灭性全面战争？我们做不到。原因不仅仅是苏联人不愿意我们知道；相反，在某种情况下，苏联人迫切希望我们了解他们的情况。问题是，他们如何才能让我们知道呢？如果中国人决心不惜一切代价统一台湾，他们如何才能让我们相信，任何制裁措施都无法阻止他们的决心，以及我们的任何威胁都将使我们卷入这场战争？[1]

[1] 缺乏有效的验证手段往往使理论处于非常尴尬的境地。在这样的博弈场景中，一方总是为对方着想。例如，一对夫妇商量是否去看电影，不管一方怎么选择，对方都会接受，双方都期望对方与自己做出同样的选择。这也是一个博弈论研究领域。这个领域比较关注双方的关系，因双方的关系和共识影响双方的价值观。我认为，邻居不喜欢我，将令我非常不舒服，反之亦然。但是，如果我们将事情公开说明白，那将令我们更加尴尬。戈夫曼指出：“社会礼节警告男士不要过早地在新年除夕约会女孩，因为过早将会给女孩留下过多寻找拒绝理由的时间。’选自 Psychiatry:Journal for the Study of International Process,18:224 (1955)。

在某些特殊情况下，博弈双方之间能够进行信息交流。在一个人为设计的博弈游戏中，纸牌或硬币代表了博弈双方的“价值观” （value system），双方可以简单地将纸牌或硬币掀起来得到结果（如果游戏规则允许或双方能够合伙骗过裁判）。但是，如果这样的行为发生在一个崇尚权力的社会，将受到严厉的惩罚，因为该社会的每个成员都坚信“要么讲真话，要么进地狱”，这一定规足以保证每个人都讲真话。然而，在某些特殊情况下，如果非要我们规定一个“范例” (general case)，那么在这个范例中，双方至少了解对方的价值观或战略选择，除非双方的价值观或战略选择不可知或不能传达。

范·纽曼 (Von Neumann) 和摩根斯顿（Morgenstern) 通过一个案例说明了他们解决非零和博弈的“方法”思路。案例主体 A 希望以高于 10 美元的价格出售自己的房子。买主 B，C 能承担的价格分别为 15 美元和 25 美元[1]（数字为作者虚构)。现在 C 打算贿赂 B，只要 B 退出这场交易，C 就可以低于 15 美元的价格买到房子。他们认为——他们的概念本身存在缺陷——B 最多能从 C 得到 15-10=5 美元。这个解决方式的条件中，有趣的部分不是 B 可能对自己的价格 15 美元做手脚，而是在实现过程中 B 根本无法进行有效沟通，即使 B 希望这么做。这个解决方式不仅——尽管考虑了所有的可能性——排除了旁观者的因素（除非他们真的希望旁观者在游戏中发挥作用），而且还假设 C 能够发现或 B 自愿暴露其真实意图。但是，这一意图连旁观者 D 和 E（被 B 空手套白狼的商业艺术所吸引的人）也无法重复。

毫无疑问，一定也存在博弈双方共同或相似价值观的博弈场景。在这种情况下，双方可以通过简单的对称原理，轻而易举地预测对方的价值观。但是在大多数情况下，双方之间是敌对关系，一方不同于另一方。即使人质小孩的父亲是绑架者，他也无法猜出自己的赎金底线。当然，倘若对方是矛矛党人（Mau Mau）或阿尔及利亚恐怖分子甚至是英国和法国的官员，也很难知道如何惩罚他。一个男孩也很难猜出心爱的

[1] J.Von Neumann and O.Morgenstern, *Theory of Games and Economic Behavior*, (Princeton, 1953), pp.564ff.

女孩对自己的评价；或者饭店的顾客也很难猜出服务员如何厌恶某种气味。

这也是为什么言语无法代替行为的重要原因。通过向双方展示某种行为的代价或风险，互动行为可以某种程度上改变游戏的进程与结局。互动行为具有言语无法相比的优势，如信息的传达、暗示方式以及不同的属性。当互动行为发挥重要作用的时候，言语往往显得异常苍白。（除非言语是一种威胁、允诺、承诺等等；而对言语的研究必须以互动行为为前提条件，而非言语沟通。）如果结果日趋明朗，互谅最终要求双方的利益分配必须根据各自的相对优势。一方的所失应该小于其所得；一方的所得应该根据某种事实依据，并非通过欺骗手段。而且双方的行为应该符合各自的价值观，能够被对方所认识。当然，这些行为应该体现言语所不具有的特征。

通常假设存在于双方价值观中的不确定因素，也削弱了作为标准原则的数学对称性方法的有效性。如果一方知道必要信息量的一半，那么数学的对称性方法就无法发挥作用；一旦对称性有助于双方进行互动行为的沟通，那么对称性将发挥质的飞跃。这时的对称性依赖于具体环境，而非抽象的价值观。

第5章

实施、沟通与战略行为

我们已经深受博弈论的影响，通常不由自主地将博弈论应用于实践中，如平时所说的威慑、核讹诈、恐怖平衡或旨在减少突袭的开放天空政策；或者我们平时形容美国驻欧洲的部队为铁丝网或厚玻璃板窗，以及认为应该让威胁体面地退出；或者我们发现威胁如此无效，以致威胁方都不愿意实施威胁。而且，我们注意到出租车司机往往拥有更大的停车位，因为人们害怕他们不在乎撞车追尾。正统的博弈理论很难解释这些行为。作者认为，非零和博弈过分强调抽象的理论研究，而忽视了这一重要的研究领域。博弈理论家通常关注从沟通和实施系统中提取理论精华，或将博弈选手的对称性看做是一个范例而非特例，从而导致博弈论忽视了一个最有可能取得丰富成果的层面，并失去了发现典型非零和博弈关键要素的宝贵机会。由于受传统非零和博弈方法的影响，博弈论以前忽视了某些典型的博弈场景或博弈模式，以及一些特殊的非零和博弈战略行为。

例如，什么“行为”能够体现有关大规模报复的争议？一个有效的威胁需要哪些要素？博弈论如何解释大家比较熟悉的“Have a bear by the tail’的故事？以及我们如何解释这个故事所暗含的收益矩阵、沟通体系和实施体系？行人如何恐吓汽车司机或小国如何对大国产生威慑？以及我们该如何使用博弈论术语表述这一行为？是什么信息或沟通结

构，或者说是动机集合导致威胁对狗、傻子、小孩、狂热者和殉教者失去作用。

某种程度上，二战后的中美苏之间动荡的冷战战略和核对峙，都可以被比作一场博弈。例如，分别处于峡谷两岸的一对敌人都在对方的毒箭有效射程之内，而且毒箭的毒性足以令对方进行二次攻击；[1]一个牧羊人将一只狼追赶到一个死角，狼只有奋力一搏，而牧羊人也不想放过狼。虽然后者随身携带了一颗手榴弹，但是由于双方距离过近而无法使用。两个邻居互相在对方的地下室埋下炸药，试图通过控制电钮或雷管来确保彼此安全。[2]如果能够分析这些博弈场景的结构或发现一些有效模式的话，我们就可以将我们的理论应用于解决现实问题。

为了进一步说明这个问题，我们这里引用一个 20 名警察围堵一名持枪歹徒的案例。本案例中，这名持枪歹徒持有一支 6发子弹的枪，可以击中 6个人。如果警方愿意付出 6 个人性命的代价的话，他们就可以制服歹徒。或者如果他们能够说服歹徒相信，对方主动缴械将免于死罪，那么他们也有可能毫发不伤地制服歹徒。当然，歹徒也有多种反抗的方式：他可以不顾警方的警告，肆意开枪或者做出某种举动显示其对警方的不信任。假设双方无法进行有效的沟通——歹徒可能使用另一种语言——他们就无法说服对方缴械。当然，警方如果不能达成共识，当然也不可能对歹徒实施威胁；或者歹徒威胁警方谁交头接耳，就打死谁，那么他就可以阻止警方达成共识。如果警方无法找到化解风险的方法，那就没有人愿意主动执行实施威胁的任务，因为他们无法令自己的威胁具有威慑力；反之，如果歹徒宣称，谁动就打死谁，那么他就可能威慑警方，除非警方有 20人一起行动，没有“第一人”之说。假设警方中有 14人的权力大于剩余的 6个人，前者可以命令后者行动，那么他们就能向歹徒证明自己能够将其制服。这样的话，他们的威胁就取得

[1] 请参阅：C.W.Sherwin，“Securing Peace Through Military Technology，”Bulletin of the Atomic Scientists，12：159~164（May 1956）。

[2] 请参阅：（Herman Kahn and Erwin Mann，“Game Theory”，The RAND Corporation，paper P-1166（Santa Monica，1957），pp.55ff.）两位作者在书中详细论述了本书中提到的炸药、威慑等方面的案例。

了成功，歹徒只能缴械投降。即便如此，他们是冒着生命危险。我们还可以假设另外一种场景。警方将歹徒逼人了死路，歹徒除了缴械别无选择，那么警方免除惩罚对方的允诺就可能发挥作用。但是，他们不能保证以后不对歹徒进行审讯。在这种情况下，歹徒最好挟持一名人质。对警方而言，他们需要彼此间达成拯救人质的默契。一旦我们能够分析得到这类博弈游戏的关键因素，我们就可以更好地理解一个暴君或少数人暴政的权力基础，或者成功暴动的必要条件。

本章将讨论博弈论涉及的几种重要行为模式和结构要素。这些行为包括“威胁”、“允诺”、“沟通破坏”、“授权”等等。此外，结构要素主要涉及沟通和实施条件等。

行　为

最典型的“行为”模式要数第 3 章中详细讨论过的承诺。如果制度环境规定，潜在买方只能进行一次发盘（相当于承诺），否则将受到相应的惩罚，那么，卖方只有惟一一次选择的机会：接受或者拒绝买方的发盘。承诺的存在将动态的谈判过程转变为具有两个极端的博弈：一方只有一次机会提出条件，而另一方只有一次选择机会。这样的话，博弈结果将更加明确。[1]

第 3 章中详细论述过的“行为”在这里将作为一个典型的行为模式。正如第 3 章所论证的，这种行为的存在和有效性取决于博弈的沟通结构，以及博弈双方是否能够找到做出承诺的有效方式及其履行承诺的能力。此外，我们假设博弈具有非对称性，“胜利”属于能够做出承诺或者率先做出承诺的一方。（我们可以考虑博弈双方特殊的关系，但是我们没有通过假设对称性的存在把双方的关系过于结论化。）

虽然为了更容易找到博弈的方法，我们可以将博弈结果明确化，但

[1] 在前面提到的卖房案例中，如果 B 找到某种方式以 20 美元购买房子并保留或摧毁房子，那么，他可以向 C 提出索要的标准，除非得到大部分，即 20–P（P 为 C 能够支付的最高价）。实际上，B 改变了自己的最高价格和向 C 提出的要求。D 和 E 也决定纷纷效仿。率先做出承诺的一方将获胜。如果 D 客观地愿意付出 22 美元，他的价格将是房子的真实价格， D 的真实性甚至大于 B。

是一旦我们确定博弈双方中谁第一个做出承诺，那么这场博弈就演化为战略博弈。尽管第一个做出承诺的一方获胜，但是这并不意味博弈如同赛跑，谁跑得快谁就赢。根据博弈规则，二者区别在于，做出承诺的一方并不意味着就一定是赢家，不管是从实际还是从法律角度，结果依赖于与第一方没有隶属关系的第二方的选择。承诺就是诱使对方采取对自己有利的战略行为，通过影响对方的预期判断而限制对方的选择。

在这样的博弈游戏中，采取承诺就如同采取了第一个“行为”。如果机制没有为双方提供禁止更改承诺的司法和契约氛围，那么博弈双方可以通过限制自己的选择自由而实现行为的不可更改。这样，鉴于一方已经做出了承诺选择，则另一方也就不可能再做出其他选择；如果他违背了这一原则，那就等于自寻麻烦。卢斯和雷夫约定，双方可以采取同样的策略应付对方，如果彼此愿意的话。例如，双方参加同一个会议这并不意味着彼此信任。“他表明了自己的立场，并保证自己不会打破会议进程，结果他丧失了其他选择权力以及届时根据自己的爱好做出行为选择的权力”。[1]同样的道理，当一个人希望在整个野外旅行过程中远离香烟时，他只能采取实际行动而非仅仅口头上的誓言。

威　胁

威胁最显著的特征是威胁方对外声称，如果被威胁方采取（或不采取）某个行为，他将采取某种威胁行为。如同承诺，一旦博弈一方做出威胁承诺，就意味着他失去做出其他选择的权力。如果威胁战术失败，威胁方面临的境遇可能会比以前更糟糕。威胁和承诺产生的共同原因在于，一方期望通过这种方式迫使理性的对方产生一种错觉，即自己可能已经改变了动机结构。和普通的承诺一样，只有当一方将威胁的信息或至少某些威胁的迹象传达给对方后，威胁才能发挥预期作用。例如，我威胁你把窗户关上，否则我就打你。你知道除非我以百分之百的决心履

[1] 参阅：Games and Decisions，pl75。

行自己的威胁承诺，否则我是不会这么做的。[1]

威胁与承诺的不同之处在于，威胁是以对方的某种行为为前提条件，承诺则是一方做出的单方面行为，而威胁是一方对另一方行为的行为反应；承诺是当承诺能够创造有利条件时，一方采取的“第一个行为”（*the first move*），而威胁则是第二个战略行为（second move）。

因此，只有当一方对另一方实施了第一个行为后，或对方实施了第一个行为后，威胁才能发挥应有的作用。当然，如果条件允许的话，假设一方必须实施第一个行为或与对方同时做出选择，他将拥有在提出威胁的同时要求对方采取“第一个行为”的“合法”权力。博弈游戏也可能存在沟通和实施的行为结构，从而为允诺提供可能性并阻止对方提前实施破坏行为。假设拦劫行人的强盗遇到了一个身无分文的行人，除非他能以人质要挟对方回去取钱，否则他将一无所获。但是这种方法必须存在一个前提条件，即罪犯为了得到钱必须在保证人质的人身安全的同时，要求回去取钱者不得报警。

实际上，类似的承诺行为或至少存在承诺特征的行为必须依赖于威胁，并要求一方能够与博弈游戏中的、与自己不同立场观念的另一方进行沟通。这一观点的核心就是，威胁必须具有必要性和可行性，而且保证被威胁者的行为对威胁客体本身的损害远远大于实施威胁的一方。卢斯和雷夫也持有同样的观点，他们以一句话形容了威胁：“伤你伤得比我深。”这充分表明，威胁取决于博弈双方的联系密切程度。如果谈判双方都试图实施有效的威胁，那么这将使扑朔迷离的谈

[1]简单地说，威胁就是当一方无法按照对方的意愿采取某种行为时，后者仅仅警告或提醒前者将对其实施报复。很清楚，一方存在实施威胁的意图。同样的场景还有当某人闯进一个人家里时，后者威胁报警，而威胁用枪将其打死则不属于我们讨论的威胁。我认为最好换一个词，如“警告”能更准确地描述这些场景。因为威胁要么显得有些花哨，不能使人想到行为；要么就是传递真实的意图或与信息结构或沟通结构相联系。在后一种情况下，威胁成为两全其美的好事，只要对方能够理解威胁的真实意图。区别警告和威胁相似性的关键在于二者是否能够准确地将信息传达给对方或让对方知道威胁的某种迹象或意图。实际上，如果威胁本质上是双方在沟通过程中出现的承诺行为，那么假设承诺出现在威胁之前，那么威胁行为的第一个行为改变了真实动机结构，而第二个起到了警告的作用。

判结果更加依赖于双方的“谈判风格”（Bargaining personality）。“在没有对对手的心理和经济状况进行全面分析的情况下，就妄下论断将是多么愚蠢。”[1]

	I	II
i	1 2	1 2
ii	0 0	0 0

	I	II
i	10 9	9 10
ii	8 0	0 8

图 5-1

[1] pp.110~111，119~120，143~144.将博弈论应用到国际关系领域的莫顿·A.开普兰（Morton A.Kaplan）认为：“任何注重双方间威胁的标准都涉及二者的效用比。”（参见 *System and Process in International Politics* [New York，1957]）。卢斯和雷夫决定他们二人之中只能有一方拥有威胁权，双方的选择矩阵为 2×2。实际上，双方不可能在这样的矩阵中做出威胁。威胁是有条件的第二次行为承诺。只有当一方希望对方做出一个选择或产生更好的效果时才能使用威胁。（如果第二个行为非常有效，那么就没有必要进行威胁，只需继续按照战略进行谈判，不必做出条件性选择。）而且在矩阵 2×2 中，条件只有对一方有利。事实上，二者进行的谈判表明选手 2根本不存在可行性的威胁选择。原因不在于选手 2 的得失大于选手 1，而在于选手 2 根本没有必要进行威胁。无论谁先选择，选择 2 一定会赢。惟一的可能性就是他要遏制对方的威胁行为，为此，他只需承诺无条件贯彻自己的战略——这也是在对方采取威慑之前的“正当防卫”。纳什将这一威胁用于谈判博弈中，并取得巨大成果，当然，他所指的威胁含义不同于这里的威胁。为了避免两败俱伤，一方不需要某个特定的结果，只需要某些有效的结果。纳什的威胁是为了取得某个特定的结果；而这时的威胁属于不对称情景。游戏规则当然允许双方改变自己的承诺，否则第一个承诺就成为最终结果了。双方可以采取合法的不作为，与对方公开谈判达成任何结果，从而逃避自己的承诺，因为只有可更改的承诺才能影响谈判过程。卢斯和雷夫的案例中非对称性是特定法律体系的一个特征。实际上，如果出现毁约的现象，毁约一方将受到社会的谴责，受到应有惩罚。不仅要对协议当事人的损失负责，而且还要对其社会损失负责。

但是，这个问题如以下面方式表述，将变得更加简单易懂。在图 5-1 中，左边矩阵的列（选择列的一方，以下简称列）将实施“第一个行为”。在没有威胁存在的情况下，列很容易取得胜利。他选择了战略Ⅰ，从而迫使行（选择行的一方，以下简称行）从 1 或 0 中做出选择。行选择了战略ⅰ，列将得到结果 2。现在我们假设行威胁列，除非列选择Ⅱ，否则行将选择ⅱ。也就是说，行迫使列在ⅱ和Ⅰ或ⅰ和Ⅱ中做出选择。当然，如果列首先选择Ⅰ，那么行无疑将选择ⅰ，双方都知道这一结果。假设列相信自己选择Ⅰ后，行必须选择ⅱ，那么这一战术就达到了预期目的。

当然，列对此可能相信也可能不相信。如果他相信这一点，威胁对其将没有任何效果，他将继续做出自己的最佳选择Ⅰ。如果他相信行一定会选择ⅰ、Ⅱ或Ⅰ、ⅰ，那么列更倾向于结果 1 而选择Ⅱ。如果换成其他数字，只要矩阵结构能够反映双方的选择倾向，这一原则同样适应，包括图 5-1 中的右边矩阵。

相对于左边的矩阵而言，右边的矩阵则更反映了威胁的显著特征，因为列对行非理性选择的惩罚更严重。但是，如果行掌握了充足的信息并做出正确的选择，行就不需要担心来自列的惩罚。列的意图非常明确，只需要行迫使自己从ⅱ和Ⅰ或Ⅱ和ⅰ做出选择。毫无疑问，列将按照自己的意图做出选择。假设我威胁你说，如果你不把最后一块面包片给我，我就打破脑袋将脑浆洒到你新衣服上。那么你只能选择把面包片给我，不管你认为我是否会这么做或仅仅将鸡蛋扔到你身上。[1]

问题在于我们是否承认博弈具有各种行为，即博弈双方是否可以在

[1] 爱德华·班菲尔德（Edward Banfield）援引印度西部某个地方的现象向我说明这一问题，他称该地区的人为吟游诗人。“在那里，人们雇用大批的保镖押运金条，即使如此，也不能保证一定安全。为此，他们与队伍的首领甚至的当地政府都达成协议。”

“他们的权力来自于个人魅力和坚强意志。如果是一个携带金条的人，他只需要做出某种承诺，即可受到保护。如果对方违约，他可以实施同样的威胁。如果没有被重视，他可能用匕首刺伤自己胸部。如果这种方式也不能达到目标，他就会刺伤自己的心脏，或者首先将自己孩子的头扭掉。不同的协议可能导致一方首先被自己的同伴砍头。对这一过程的恐惧和担心死人的血溅到头上可以使人们保持理性。由于不忠将要受到如此严厉的惩罚，所以他们都不惜一切代价去履行自己的承诺。”（Mountstuart Elphinstone *History of India*［ed.7;London，1839］）

博弈过程中实施某种行为，从而改变博弈进程——改变结果的选择方式、选择的顺序、博弈的信息结构。假设除了双方达成共识或拒绝对方的条件外，博弈本身不允许任何其他行为的存在，那么博弈的结果很大程度上取决于博弈双方的“个性”因素，即他们在一个没有行为的博弈过程中只能通过心理沟通达到相同的预测判断。如果威胁不是一方对另一方的强制警告，我们可以理直气壮地质问，什么是真正意义上的威胁。但是可以肯定的一点是，威胁一定涉及某种承诺因素——不管是真还是假。

这里的承诺一词含义十分广泛，既包括迫使对方必须做出选择的行为（如一方司机以加速行驶威胁另一方退缩）；也包括改变一方此前出于经济原因而做出的最终选择（当政府故意惩罚虐待狂；或投保人改变自己的要求和向保险公司做出的允诺。）；还包括为避免履行诺言造成的损失而毁约的紧急自损行为（如一方承诺如果自己不能履行协约将当众出丑；或一方在自己的商店前安装厚玻璃以警告女顾客和小孩。）在日常生活中，我们也能看到类似的例子，如欧文·戈夫曼（Erving Goffman）为我们提供的“商人”案例。戈夫曼指出：“商人，特别是街边的小贩都非常清楚，排起队伍卖东西往往缺乏信誉感，从而赶跑顾客，除非一些顾客为了卖方的面子而勉为其难地上当购买，但这也掩盖了通常发生的一幕。”[1]

然而，有时对威胁承诺的界定并非如此严格。博弈一方可能认识到“实”（firm）承诺意味着必须履行，否则将受到严重的惩罚。换句话说，一方一旦做出了这样的承诺，就必须不惜一切代价去履行。这种无限（至少是严厉）惩罚是他的自愿选择，并且不可反悔，具有可衡量性、可操作性。在整个博弈过程中，做出承诺的一方必须以实际行动履行自己的承诺。当然，条条大路通罗马。如果把惩罚的力度适当降低以适应于其他情况，也可能达到同样的效果。在图 5-2 中，如果列首先行动，那么他将获得胜利，除非行承诺选择 i（承诺为行赢得‘第一个行

[1] 戈夫曼的论文对博弈论与博弈者之间的关系进行了细致分析，发展了对标准博弈行为结构的研究，如礼仪、骑士精神、外交惯例和法律。

	Ⅰ	Ⅱ
ⅰ	4, 2	1, 1
ⅱ	4, 2	3, 3

图 5–2

为”权）。但是如果承诺意味着对选择 ⅱ一方的有限惩罚（如图所示，我们从行的选择 ⅱ 中各减去 1 以示惩罚），那么只有当惩罚大于 2 时，这时承诺才有效。对于列而言，如果他选择 Ⅱ，那么行只能选择 ⅱ 并放弃这一承诺权。因为，在这种情况下，承诺对行来说只是个负担，因此行竭力避免做出任何承诺。

与上面的情况不同的是，如果图 5–3 没有威胁的存在，不管游戏规则要求谁首先行动或是否同时行动，游戏的结果都将为 ⅲ和 Ⅱ。只要一方拥有“第二个行为”权并向对方施加威胁，他一定是赢方。[1]列可以威胁 Ⅰ–ⅲ；行可以威胁 ⅰ–Ⅱ。但是如果惩罚保证了威胁有效实施，那么列期望的较轻惩罚结果是 4；当行选择 ⅲ时，列只有选择 Ⅱ 才能得到他希望更轻的惩罚。如果行不履行承诺，那么他可能得到的最轻惩罚将是 3。假设惩罚程度具有惟一性，不能低于 3，那么结果一定是 ⅲ和 Ⅱ。如果惩罚大于 4，那么率先做出威胁承诺的一方将获胜。当收益大于 3

[1] 例如，如果选手列不能按照规定迫使选手行首先选择，那么列可以“合法地”威胁行允诺选择 ⅱ，自己选择 Ⅰ。分析过程中，我们应该注意对允诺和威胁承诺的惩罚力度。尽管对允诺的人为或机制（对第二方而言）本质上不同于单方面允诺（第二方不能自主解除）。一般而言，对违背允诺和威胁承诺的惩罚不同，这与对双方的惩罚是同样的道理。图 2–4 的结果就必要惩罚条件必须大于或等于 1，无论是列还是行做出允诺。请注意，如果威胁以允诺为条件，那么双方存在违约的情况，则实施威胁的一方将获得优势，而被威胁一方将处于不利地位。

i	-5, -5	-2, -1	-2, -1
ii	-4, -3	0, 3	2, 2
iii	-5, -5	1, 1	3, 0

图 5-3

小于 4 时，惩罚仅对赢方行有效。在后一种情况中，当威胁失败时，损失最大的一方就是不能实施威胁的一方——但是他不能进行严厉的自我惩罚。

这里需要注意的是，“更大损失”（hurt more）并不是指行的威胁造成双方谁的损失更大，而是指行完全履行威胁承诺时是否会比列遭受更大的损失。这里不是行实施威胁行为，而是列。实际上，如收益矩阵所示，只有当履行威胁承诺遭受的损失比列大时，行的威胁才能取得成功；反之，只有当履行威胁遭受的损失比行小时，列的威胁才是成功的威胁。

此外，扩大威胁含义的另一种方式是将我们的假设理性化。假设 R 为 Pr，C 为 Pc，而且双方有可能做出错误、非理性的行为，或者由于对另一方选择结果的误解而盲目地决策。[1]现在我们得到一个不同的博弈游戏。在这个游戏中，双方履行威胁的得失涉及到一方威胁的完全履行是否引起对方的注意。如果一方因实施威胁而遭受的潜在损失大于对方，那么这个博弈游戏就具有对称性——双方拥有平等的权力，可能遭受同样的结果。在这个博弈环境中，一方可能发现威胁对自己有利，对对方不利；而对方可能持有不同的观点，坚持威胁的错误性。（同样的道理，如果双方拥有实施威胁的同样机会或双方能够同时实施威胁，而

[1] 第 7章和第 9 章将对这一问题进行深入探讨。

后一种情况往往是由于一方没有认识到对方的承诺，而无法及时停止实施威胁。）对威胁概念的发展——在理性条件下——很大程度上是以“更大损失”为标准。总而言之，博弈论往往强调，如果博弈一方实施威胁而非仅仅局限于口头宣言，那么，其收益必须大于对方，其损失必须小于对方。这无疑是对谈判战略理论的巨大贡献。除非排除相对“效用优势”（utility-comparison）标准，否则我们根本无法认识和分析诸如战争威胁、价格战威胁、文明社会对犯罪行为和行为不端进行惩罚威胁以及勒索、敲诈和威慑。博弈双方之间的威胁不对称性还需要我们进行深入的研究和探讨。但是非对称性仅是就某些方面而言，如沟通系统、威胁和允诺的可行性、做出承诺的时间、预期反应的理性化程度以及相对损失标准。

允　诺

我们不能把强制性的允诺认为是理所当然的事。一份有效的协议必须以强制性术语表达，并涉及协议当事人的强制性行为。强制性取决于至少两个因素——拥有相当权力的权威部门能够判决是否实施惩罚和拥有相当权力的权威部门能对双方的不当行为进行必要的惩罚。战后不计其数的裁军谈判和核查计划已经表明这一问题何等艰巨，即使谈判双方应该达成一个强制性的协议或有效的强制实施方案。如果谈判双方没有互信或根本不相信对方有任何履行协议诚意，这一问题将变得更加复杂。如果存在有效的方式迫使双方做出强制性允诺，或世界各国都能毫无保留地承诺某种世界权威，那么武器核查的技术问题就会消失殆尽。但是由于没有有效的监督手段，甚至有效的惩罚机制也无法迫使各国做出强制性允诺。当然，有效惩罚机制的匮乏无疑加剧了这一问题的严重性，除非公平谴责最初协议的第三方愿意出面执行这一惩罚。此外，我们必须排除掉一些看似重要却不具有操作性的协议。通常只有内容客观公正、拥有必要监督机制的协约才具有可操作性。

允诺必须是对等的双边（契约式）承诺，双方的允诺互为条件。当然，在某种情况下，为了促使对方能够从双赢的角度做出选择，一方可能单方面做出允诺。图 5-4 就说明了这一点。在左边的矩阵中，如果双

	I	II
i	0, 0	2, -1
ii	-1, 2	1, 1

	I	II
i	0, 0	2, -1
ii	0, 0	1, 1

图 5–4

方同时做出选择，那么只有一次允诺是有效的；在右边的矩阵中，如果列一定选择 II ，从而出现双赢的局面，那么，行的允诺将为其赢得收益。（如果在左边的矩阵中，双方轮番做出选择，那么，第二个选择的一方将拥有做出允诺的权力。如果双方约定选择的先后顺序以及其中一人将做出允诺，那么，他们可以决定另一方首先做出选择。与右边的矩阵中单方面无条件允诺不同的是，这些允诺必须以第二人的选择为前提条件。）在前一种情况中，如果一方能够将对方消灭掉而不被觉察，那么，他就会破坏这一允诺[1]。如果一个国家即将拥有完全的突袭能力，那么，它可能率先单方面做出允诺——如果可能的话——从而避免来自敌国先发制人的袭击，因为对后者而言，这是惟一的机会。

允诺的真正含义——例如威胁—— 非常模糊。允诺似乎是第二方非常欢迎的允诺（有条件或无条件）。在图 5–4 的两个博弈游戏中，允诺都是双赢的选择。但是在图 5–5 中，行必须在选择威胁的同时选择允诺。行不仅要威胁 ii – I ，而且还要在对方选择 II 时允诺选择 i 。行 的这一选择将使列的收益是 4 而非 0；如果后者选择 II ，这是非常有利的

[1] 约翰·科利尔的《多雨的星期六》反映了这一思想。最近还被艾尔弗雷德·希区柯克制作成电视片。一个人无意间目睹了一场谋杀案。凶手为了堵住他的嘴，迫使他在尸体上留下手印和其他犯罪证据。这样一旦事情败露，他也将受到法律的严惩。实际上，他应该制造假证据，与真正的凶手一起承担罪责。〔摘选自短篇小说 *New Yorker* [London，1951]，pp171~178)]。

	I	Ⅱ
i	5 2	4 4
ii	1 1	0 5

图 5-5

收益。如果行无法做出允诺，那么列将赢得 5。因为没有允诺，威胁无法发挥应有的作用，甚至根本不会出现威胁。行的威胁 ii－I 对其一点好处都没有，这一选择无法迫使列选择 Ⅱ，因为如果列做出了这个选择，他只能得到 ii－Ⅱ，即 0 而非 1。只有行做出允诺，他的威胁才能产生威力，允诺的使命就是为威胁服务，迫使列得到 4 而非 5，为自己赢得 5 而非 2。一方显然不能仅仅通过不遗余力地追捕而迫使间谍、阴谋者和病毒携带者自报门户。当然也不能豁免自投罗网的人。[1]

一种更好的解释或许是，允诺是第二方可以控制的承诺。换句话说，第二方可以自主地选择履行与否，但是必须在有效的时间内。只有一方完全做出威胁承诺后，允诺才能生效。但是，如果允诺的对方（列）事先做出拒绝声明，那么行知道自己选择 Ⅱ 后，对方会选择 0，则威胁将被化解。如果威胁允许不可分割或带有不能撤回的行为，那么这个定义将变得十分模糊。（事实上，一旦某种不可更改的行为而非法律承诺可以取代允诺，这种定义将无法成立。）

实际上，只要存在两个以上的选择，威胁和允诺就成为一方对另一方的互动反应模式。鉴于此，尽管威胁和允诺都属于选择性和条件性的自我承诺，我们最好将二者看做同一战略的两个不同方面。有些案例已经表明，这种承诺很大程度上考虑了对方的利益。

[1] 某些自首者可以从轻处理。

实施方案　如果不存在权威的监督或有效的监督方式，那么协议根本没有强制力。实际上，协议通常涉及的问题主要是协议的形式和协议内容的协商。因为一种有效的方式和明确的内容能够阻止令双方不愉快的事情发生，如一方采取欺诈行为，则保证协议的有效实施。在强调双方之间互信重要性的同时，我们还不能过分依赖这种互信，因为甚至是博弈论对互信的定论本身就值得商榷。互信通常依赖于双方之间关系的稳定和双方对长远利益与短期利益的权衡。如果对方认为，长期利益重于短期利益，那么，他就不会为了短期利益而采取欺诈行为；反之亦然。当然，如果双方间的互信发展到一定程度而发生质的变化，那么，在旁观者看来，互信好像是某个特定阶段的产物。

在以下几种情况下，人们往往能够达成有效的、具有强制力的协议。协议的成立以双方的合作与某种补偿为条件，如两个人就吃饭的地点发生分歧；两个罪犯就犯罪证据发生分歧；公司员工就报价或足球队员就战术问题出现分歧等。在这些场景中，争议双方存在某种共同利益高于二者之间的分歧。为了实现终极目标，双方需要克服分歧，达成共识。协议一旦成立，就成为双方行为的标准，为以后合作提供依据。协议当事人除了按照协议履行自己的责任，别无选择。如果双方自己无法以有效的方式监督协议的履行，他们可以寻求其他监督机制的帮助或者引入第三者监督人，后者的使命就是严厉惩罚违约一方。在生活中，我们可以看到这方面的案例，如将藏宝图一分为二或将枪支的使用权交给一方，而将子弹交由另一方持有。

古代的人质交换机制也值得博弈理论家好好研究，类似的例子还有典型的同饮一杯水以防对方下毒，或在大庭广众之下会面以防对方埋伏，以及在毒品交易中使用吸毒人员作为中间人或代理人。这些例子都是单方面人质行为的典型例子。

如果两个敌对国家想避免两败俱伤，它们可以交换人质以示诚意或达成一个共识将两个国家的管理部门迁移到某个孤岛，由两个国家轮番管理。在此前提到的双方价值观问题中，都没有涉及对互换人质和采取理性行为的重要障碍问题。一个国王将自己的女儿送到敌国作为人质，但是这样并不代表对方一定相信他的诚意，因为对方或许还认为这个国

王根本不喜欢这个女儿而这么做的。我们或许可以通过开展“海外儿童年”活动（junior year abroad）来消除前苏联人对来自美国突袭的担心。具体做法是：将每一个5岁的儿童都送到前苏联学习一年——反之亦然。这么做纯粹是为了发挥人质的作用，而非文化交流——每一届的新生都在毕业生离开之前到达，那么，美国就没有任何机会对前苏联实施核偷袭。当然，我们无法肯定前苏联人和我们有相同的想法；我们也无法肯定，这样的双边计划是否能对前苏联人发挥同样的作用。不幸的是，即使这个双边计划对前苏联人发挥了同样的效果，但是我们也没有充分的理论说服自己相信。尽管如此，在许多突袭情况下，单方面的允诺总比没有要好得多。交换人质的做法值得考虑，即使不是对称性的交换。[1]

实际上，交换人质思想与认为大国进行裁军谈判更有效（或许涉及更多技术问题）的观点具有同样的逻辑，如果后者涉及国防军备及其结构的话。本质上讲，一个国家没有国防就意味着将自己的国民作为人质，区别在于并没有将他们置身于对方国界内而已。如果这样的话，前苏联人将他们的孩子交给我们，我们就可以对他们进行虐待，并置之不理，任其生灭；反之亦然。这样就现出了我们通常所说的“恐怖平衡”（Balance of terror）——实际中确实存在这样的情况——这如同人质大交换。（这一类比要求平衡比较稳定，双方都不能偷袭对方，打击其反击能力，与此同时，双方只能进行前面提到的行为。）[2]

违约行为　允诺的有效性还与第三方的影响联系密切，例如后者可以故意破坏双方间有效的协议，增加双方的困难。这意味着，禁止非法行为将是对协议双方的放纵。如果这样的话，合同也就没有权威性和有效性了。不能履行赌博、贸易制裁或在禁酒令期间走私酒品等契约，都是人们抵制这类行为的表现。当然，有些时候这类禁令可以成为某些人

[1] 要准确地定义人质确实比较困难，如同定义允诺和威胁一样。如驻欧美军可以被看做是人质，尽管他们的使命是为了表明美国可以插手欧洲冲突事件。如果他们不能发挥人质的作用，那么他们的孩子和妻子可以做到。一般而言，侵略国总是避免在被入侵国的旅游旺季入侵，因为这样可以避免不必要的人质问题。

[2] 第10章将对这一问题进行深入探讨。

保证双方履行协约或允诺的权力。[1]在禁酒令期间，不保护专利酒品意味着只有大的非法组织才能生产高质量的酒品，禁酒令无疑是在帮助他们垄断市场。同样的道理，保护专利权和商标的法律可被看做是促进非契约贸易的手段。

主动权的放弃

是什么造成威胁和普通的承诺成为难以实施的策略？一个有趣的研究课题是，找到履行承诺和惩罚违约行为的有效方式。限制或放弃一方选择权的相关策略多种多样，数不胜数。这些策略只有一个目的，摆脱令人尴尬的主动权，从而使结果取决于另一方的选择。

这就是国务卿杜勒斯希望在下面的文章中寻找的策略。

在不久的将来，我们很有可能摆脱对大规模报复威慑的依赖……因此，与 20 世纪 50 年代不同的是，在 20 世纪 60 年代，那些以中国和前苏联为核心的国家能够具备有效的国防力量阻止全面的常规进攻。面临敌国入侵时，他们也有权力在投降和使用核武器与敌人同归于尽之间做出抉择。从这个意义上，我们可以说格局将发生变化，问题不是被侵略国是否具备核打击能力保护自己以及侵略者不应该具有常规进攻能力，而是他们如何权衡核战争的后果。[2]

杜勒斯认为 20 世纪 50 年代的威慑与 20 世纪 60 年代的威慑不同之处在于，谁是最终的决策者。这一差异非常重要，因为美国还没有找到一种令自己信服的有效承诺方式以大规模报复行动对付某种进攻行为。

在第一颗原子弹成功付诸实践以后的一段时间内，有人怀疑地球的大气层对核裂变的承受能力是有限的。这一担心来自于核爆炸的连锁反

[1] 有人认为敲诈者的作用在于保障非法契约的履行。根据 R.L.达弗斯（R.L.Duffus）的观点，芝加哥服装贸易在降价行为可能遭到爆炸的报复——定价组织指使。参见 R.L.Duffus, The Function of the Racketeer”, *New Republic*（March 27, 1929）, pp.166~168)。

[2] J.F.Dulles, “Challenge and Response in U.S. Policy,” *Foreign Afairs*（October, 1957）迪安·艾奇逊在其著作 *Power and Diplomacy* [Cambridge, Mass, 1958], pp.87~88）中用同样的语言论述了其对美在欧驻军的立场：“我们应该让对方相信，任何大规模进攻都将导致大规模报复，因为‘他将为我们决策……我们在欧洲驻军为了更有效地进攻’。”

应将严重破坏地球的大气层。有人认为，如果这一怀疑成立，那么，我们可以准确地测定地球大气层的承受能力，从而可以将核弹的使用控制在极限范围内，即 n–1 个核弹。

美国空军中校（后降为少校）史蒂文森·坎农成功演绎了围魏救赵的战术。在他的条幅式侦察照片中，坎农驾驶自己的飞机保护一艘即将被解放军地面部队捕获的国民党舰艇。在没有敌意和攻击命令的情况下，尽管知道没有必要实施威胁，他还是驾驶飞机将一个燃烧的汽油桶投向地面部队，从而迫使解放军退让。他既不需要向敌船投掷汽油桶，也不需要实施威胁，由此他放弃了主动权。

同样的战术还有将“消极抵抗”称为“积极不抵抗”。根据《纽约时报》（The New York Times）的报道：“今天，日本罢工的铁路工人在 300 多个火车站的轨道上示威，造成 48 人滞留，144 个班次取消。”[1]

《纽约时报》还报道了一个类似的日本新闻。本周日本进行了一次公开辩论。辩论的主题是，是否允许自杀式静坐抗议船驶向复活岛（Christmas Island）周围的禁区水域，英国计划在该岛进行氢弹实验……这次远征的首要目标就是阻止英国的核实验。[2]

定　位

任何博弈游戏都涉及一个重要问题：博弈双方对彼此的价值观究竟知道多少。还有一个问题涉及双方完全定位的问题。如果一个欲打劫银行的银行职员找到一个非银行内部人员作为同谋者，他也许会付诸行动；但是，如果他只能找一个银行内部人员为同谋者，那么，他可能并不付诸实施。因为二者同为银行职员，谁也不难确认对方是否真得敢于打劫。谁也不会将自己的真实意图告诉一个可能没有共同利益的人，否

[1] “Rail Strikers Sit in Tracks,” *The New York Times*（*May 13，1957*）*.Pp14 L f.*）“火车司机将火车缓缓行驶，故意从司机座位上跳下火车，穿过车站后又跳回驶过来的火车。他的立场弱点是火车在工人离开轨道之前即能停止，特别是罢工工人故意集中阻止火车。他们可以通过占领轨道并将司机的钥匙丢掉来阻止司机，如果他们能在司机离开火车之前警告他的话。”

[2] “Japan Debating Atomic ’ Suicide,” *The New York Times*（March 5，1957），p. 16）.

则一旦事情败露将受到严厉的惩罚。类似的情况，一个男孩不敢邀请自己心爱的女孩约会，惟一的原因是他害怕对方拒绝自己。同样，一个绑架者如果不能将穷人和富人区别开，他也不能实现自己的目标。美国南部反对种族隔离的少数人，可能永远不知道对肆意散布反动言论的惩罚力度。

和沟通一样，定位也不一定是互动行为，一方有时可以随意地改变对自己的定位。一旦确定自己利益实体，谈判主体可能在谈判开始之前就已经对相关问题进行了准确定位。莎士比亚的《针锋相对》（Measure for Mearsure）充分体现了这一原则。主人公安吉洛（Angelo）是一位公爵，他关押了一名罪犯，并希望将其处死。当然，他也可以让其生不如死，但是他没有这样做。罪犯的姐姐前来为弟弟求情，希望公爵饶其一命。公爵看到罪犯的姐姐颇有姿色，顿生非分之念，并提出过分条件，但是被后者断然拒绝。公爵就威胁罪犯的姐姐，如果她不答应条件，就折磨其弟弟。现在由于双方自身的定位和沟通的实现，一场博弈场景在他们之间展开。很显然，安吉洛的真实意图并非要折磨罪犯，这只是他的一个要挟条件。如果有一个人出来接受这个条件，显然安吉洛的这一条件将为其带来收益——而不是折磨罪犯本身。现在，罪犯的姐姐处于不利境地，不仅确定了自己的身份，而且收到了对方的要挟条件，那么，她现在必须承受她不应该承受的痛苦。因为她在不暴露自己的身份或在对方提出条件之前逃走的话，就可以避免这一遭遇。

几年前在纽约郊区发生了一件博弈案例。几个摩托车飞车族身上带有会员证，以向警方证明自己是某个摩托俱乐部的成员。如果他们被警方逮住，则可以向警方证明自己是俱乐部成员，并向警方贿赂。这些会员证的作用就是向警方证明自己的身份，并且提示警察（如果后者接受贿赂的话）保持沉默。它也证实骑摩托者的允诺是有效的，但是只能在他们被警方逮住之后。假设警察认识他们，而且仅仅是惩罚驾车的摩托车手，并威胁他们交罚款，那么，现在这些会员证能否发挥作用则取决于摩托车手是否使用它们。萨瑟兰（Sutherland）描述了一个类似的场景："大多数警察都会或多或少地公平处理各种犯罪案件，因为他们拿的就是这份钱。但是，他们有时也在接受贿赂后对自己的罪犯网开一

面，尤其是一些非惯犯。因为他们知道这样做非常安全，上级不会知道．特别是对非惯犯。”[1]

定位还与一个重要的经济现象关系密切，而传统生产与交换经济学往往忽视了这一现象，即可能存在大量的破坏因素并与其产生的威胁有关。例如，据作者粗略估计，一个智力略低于一般水平的普通大学毕业生一年辛辛苦苦创造的价值也许只有 3000 美元或 4000 美元；但是如果他愿意的话，可以破坏掉比这一数字大百倍的价值。如果存在一种机制，小伙子能够从其破坏的价值中获益匪浅，那么，他很有可能选择放弃技工或办事员的职业，而成为勒索者，专门从事破坏活动。令人欣慰的是，敲诈很大程度依赖于自我定位和勒索犯的自我沟通。

通过安排犯罪嫌疑人与受害者接触，可以有效地增强罪犯的自我定位。同样的道理，如果证人受到某种威胁时，为了保证在罪犯被捕之前不泄露目击证人的身份，通常在大陪审团面前进行秘密宣誓。（博弈可以被广泛应用于法律政策和法律的实施以及刑事威慑领域。）

授　权

博弈一方有时将自己决策的部分主动权以及全部或部分利益授权给可能成为（或已经是）对手的委托人。实际上，这一“行为”时常发生。由于保险计划允许利益共享，因为保险公司通常都准备不同的方案，从而能够更好地威慑投保人或应付投票人的不轨行为。同样，出于互相制约的需要，有时我们需要不同部门的签署批准。雇用催债公司替自己讨债就是为了实现与债务人的单方面接触，从而有效地避免来自债务人的威胁或借口。向韩国军队提供军火或允许他们接受战俘营是摆脱令人进退两难的权力的一个策略。之所以令人进退两难主要指要么自己不得不卷入强制性或威慑性威胁，要么迫使自己从自己的威胁中让步，这无疑削弱了威胁效果。

美国与台湾的《共同防御条约》可以被看做是向意志坚决的一方移交权力的一种形式。近来有关允许欧洲国家拥有核武器的建议引起了很

[1] E.H.Sutherland，the Professional Thief，Chicago，1954，p.126.

大争议。有人认为，将一种显性权力交给在某种情况下比美国意志薄弱的国家，将增大威慑存在的可能性。

雇用暴徒和虐待狂收取保护费或充当罪犯保镖，或是将权力移交给一个军事指挥官都可以被看做是金蝉脱壳策略的应用。因为一旦威胁失败，幕后决策者可以轻松地推脱责任。（这正如一个理性博弈选手理性地装疯卖傻一样，从而避免外界威胁或推脱实施威胁的责任。同样，一个理性的博弈选手也可以选择一个非理性的同伴或委托人。）

图 5-6 中（忽略括号内的数字），如果行是第二个选择的一方，那么他将失去右角的收益，列得到有利的收益。作为一个副产品（括号内的结果），如果存在一个没有决策权的第三者，只要行将其选择权移交给第三者，那么行就能取胜；后者也将得到左上角的收益，而非前面行得到的收益。尽管行不得不将自己的一部分利益分配给第三者，但是他还是有必要支持第三者。对行来说，这是利大于弊，因为这样的话，他将得到左上角的 3 而非右下角的 1。

	I	II
i	3 (2) 2	2 (1) 0
ii	4 (1)	.5 (0)

图 5-6

调　停

调停者的作用也是博弈论研究的一个重要课题。无论是博弈游戏规则的自身安排，还是选手为了尽快达成共识而自我约定，调解人可以被看做沟通机制的重要因素或能够影响沟通，并取得其自身利益的第三者。但是，调解人的作用不仅仅在于影响双方的沟通——规定双方的递盘规则等等——他还可以制定自己的规则，并提出积极的建议。也就是说，在双方选手无法沟通的条件下，调解人可以按照自己的意图来影响

双方的预期判断。一个行人也可能临时指挥交通，疏缓堵车，这是因为堵车的司机为了尽快摆脱堵车而愿意将指挥权交给他。虽然指挥权只是建议性的，但是协作需要堵车司机共同承认这种建议权。还有一种情况，或许每一对舞伴都愿意跳自己喜欢的曲子，但是只要主持人拥有一个麦克风，就能迫使他们接受自己的曲子跳舞，不管他们喜欢与否。通常情况下，路中间的白线就是一个调停人，只要它能公平地对待马路两边的行人，并赢得他们的承认。有关夏令时的争议也说明这一问题。如果大部分人希望早一个小时工作，那么他们必须能够控制时钟。当然，即使他们这样做了，那些持反对意见、组织良好的一小部分人，往往不会改变自己的作息时间来适应这一调整。

调停人也可以帮助博弈双方保留某些积极条件。一个调停人可以发挥某种沟通的作用，也可以阻止某些信息的传递。（从这个意义上讲，调停人可以是一台计算仪器。）例如，调停人可以通过比较双方的发盘，然后指出双方存在的差距，但是不用暴露实际的发盘。他也可以是一部扫描仪，将一部分信息压缩其中。但是，调停人只能做出某些博弈双方无法实现的有限比较，因为没有选手专心致志于忘记什么事。

还有一个问题就是，如何解决左手得到右手需要的信息的问题。例如，政府的一些部门需要准确的收入资料用于数据统计分析；但是，政府的另一些部门可能需要这些数据用于增加税收或审查偷税漏税。于是，为了得到第一手资料，有关政府部门就有必要寻找一种方式保证这些资料不被落入有关税务部门手里。类似的情况，一些公司调停人通常只有在某种条件下才能将公司的商业秘密交给统计部门，如税务部门必须保证资料的用途有利于公司并保证使用完后将资料销毁；咨询调查机构必须保证以集体的方式发布信息，不能直接涉及个人隐私。当大宗货物的顾客认为如果物品所有者不知道谁是真实的买主，其即可支付便宜的价格时，调停人就发挥阻止双方身份定位的作用。

如果博弈双方决定将不可更改的权力交给调停人，那么调停人将发挥仲裁者的作用。但是只有当选手双方遭受损失时，允许仲裁者惩罚或给予其某种补偿的时候，仲裁协议才能发挥作用。当然，博弈双方必须相信仲裁者，并对其做出强制允诺。但是，如果仲裁者拥有强制双方履

行允诺的全部手段，那么一旦双方不信任对方，他们可以找到双方都信任的第三者保管其利益。[1]

沟通及其破坏性

许多有趣的博弈游戏和博弈场景都依赖于沟通结构，特别是沟通的非对称性和一方对初期沟通的单方面选择或破坏。如果双方无法进行有效的沟通，那么，即使是威慑也无法产生预期效果。勒索场景更是如此。如果勒索人不给对方选择的机会，那他根本无法达到目的。更有趣的是，假设你对哭着的孩子说“不要哭，我就给你好东西”，但是孩子哭得非常凶，因为哭声掩过你的声音而没有听到你的话，那你还能达到目标吗？（有时孩子对此心知肚明）如果一个目击证人因受到监护无法与外界沟通，那么他就不可能出于对恐吓的担心而作假证，即使他可能已经推断自己不可能受到威胁。

如果博弈结果取决于双方的合作，那么一方如果能及时破坏沟通可能也不失为一个致胜良策。假设一对夫妇为吃饭的地方而在电话中发生争执，这时，如果妻子简单地告诉对方自己要去的地方，然后把电话挂断，那她就是赢者。同样的道理，如果一个人不想就选择的问题与对方争执，他可以简单地将助听器摘掉。

正如本章前面所讨论的，暴乱行为的成败往往取决于沟通是否顺畅。如果沟通不畅，那么当局就可能通过切断暴乱分子的沟通渠道，从而防止暴乱的出现；反之，如果暴乱分子能够有效地沟通，并建立统一的识别系统，那么，他们就可以令当局闻风丧胆。假设暴民熟悉警方或平时双方就是邻居，那么，暴民进行暴力活动等潜在威胁就可能被警方知道。那样的话，适当借用外力可以有效地化解暴民对当局造成的威胁，一部分原因是他们削弱了威胁的发生，主要另一部分原因是切断了暴民之间的沟通渠道。联邦部队之所以成功地在小石城平息威胁活动，

[1] 有人告诉我在没有强烈商业诚信传统的国家里，为了公正和公平，交易双方往往邀请具有诚信文化背景的商业伙伴或董事作为监督人。因为在后者看来，诚信与公平是他们的共性，具有崇高的价值。

很大程度上是因为与当地的警方相比，他们与当地民众之间没有过多的沟通和共同的价值体系。州部队之所以能够平息 1943 年底特律种族骚乱事件，而当地警察没有有效地发挥作用，也是同样的原因。使用摩尔人和印度锡克教徒以及其他语言的部队成功镇压当地的起义，很大程度上取决于他们与暴民没有共同语言，无法进行沟通，从而无法得知对方或受害人可能传达的允诺或威胁。甚至将军官与士兵分开也是为了避免军官收到或觉察到威胁，从而不能有效地抵制和化解威胁。

当然，对威胁方而言，是否知道对方收到了自己发出的威胁也具有重要的意义。如果他主观认为什么能够传达给对方，但实际情况却没有，那么即使他发出威胁信息，也无法达到目标，结果可能不得不进行一次毫无意义、两败俱伤的威胁行动。因此，镇压暴动的士兵必须来自外地并不断地转移，以避免与当地民众“熟悉”而且还要保证他们免受镇压对象的影响。例如，避免引起镇压对象的某些注意；避免因为对方的讥笑而害羞；避免认识镇压对象，即使有人故意鹤立鸡群。在实际运用中，防暴士兵最好面戴头盔，身穿制服，消除他们与暴民沟通的可能性。

事实传递　沟通不仅仅指讯息的传递，也指一些事实的传递。假设一方想对对方进行威胁，那么，必须将威胁承诺传递给对方，允诺亦然。而承诺的传递不仅仅是言语上的传递，还涉及有关承诺的各种事实(evidence)。换句话说，一方若想将威胁传递给对方，他就必须让对方亲眼看到某些相关的事实或找到一种能够辨明是非的仪器。如一方可以通过邮寄将一张签字支票传递给对方，但是他却无法通过电话向对方证明自己已经签字，因为对方在没有看到真实支票之前是不会相信他的。同样，一个人可以向别人说自己的枪已经子弹上膛，但是如果他没有实际行动，那就没有人相信他所说的话。从博弈论的角度讲，上面的区别就如同巴黎传递信件与普通电报系统的区别；电视与收音机的差异。（中间人的作用之一就是判断一方向另一方所说之言的真实性。）例如，密码系统使人们通过电话使周转现金成为可能，来自银行密码系统的证实使收款一方确认交易成功。艾森豪威尔的“开放天空”(open-skies)政策验证了事实传递的重要性和困难度。类似的情况还有一些旨在消除

双方对偷袭的担心的一些机制。利奥·西拉德（Leo Szilard）指出了一个看似自相矛盾的现象：一方希望将豁免用于敌方的间谍而非对其进行折磨，因为只有间谍能够说服对方相信[1]自己没有任何偷袭的意图。

有趣的是，政治民主本身依赖于事实沟通匮乏这一博弈结构。无记名投票难道不是为了剥夺投票人出卖选票的权力？这不仅仅是简单的保密问题，而且是强制保密，是无记名投票机制运行的前提条件。投票人必须放弃其他投票方式。投票人失去的不是可以出卖的财产，而是受恐吓的权力。现在，他再也不能收到恐吓信了。假设投票人可以任意出售自己的投票权，那么很难想像有多少人对他实施威胁；除非他出售自己的投票权，否则没有人会罢休。但是，现在投票人无法向别人证实自己是否服从威胁，那么他和威胁他的人都知道任何报复都与他的实际投票方式无关。这样的话，任何威胁都将是徒劳无功、白费功夫。

摩托手与交警的故事也是有关默式、非对称性沟通问题的典型案例。一个摩托手在十字路口遇上了正在指挥交通的交警。假设前者真看到了交通警察的指挥信号，但故意视而不见，那么，交通警察就会给他一张罚单。假设前者假装没有看到交通警察的指挥信号，并违反交通规则，那么，交通警察可能仅仅认为他的行为是非故意的，而不会对其实施处罚。假设将摩托手换成汽车司机，如果后者知道交通法律，但故意违反交通规则，那么，是否处罚则掌握在交通警察手里。交通警察可能会放其一马，也可能出于警察的威严而对其实施处罚。小孩通常对此谙熟于心，经常利用这一规则逃避父母的惩罚。孩子经常故意回避父母警告的眼神，因为他们知道一旦做错了事如果看到了父母的眼神，就可能受到父母的惩罚。与孩子相比，大人并不逊色。他们也会利用这一规则避免向上级提出可能不被批准的要求，因为他们知道先斩后奏更有效。[2]

[1] L.Szilard，Disarmament and the Problem of Peace，*Bulletin of the Atomic Scientists*，2：297~307（October，1955）.

[2] 戈夫曼（Goffman）赋于沟通以“法律地位”：“具体操作上，交易双方通常通过言语暗示——影射、模棱两可、适当的停顿和幽默的语言——实现商务合作。这种非正规沟通技巧要求双方完全不能故装正襟严言。但是，对方却有权力装疯卖傻，故装不知。可以说，暗示

沟通结构的有效性取决于博弈双方的理性程度。“Have a Bear By the tail”的故事充分说明了这一观点。双方达成结果的最低要求是，熊必须有能力做出有效的允诺并拿出具有说服力的保证，如一旦违约，熊将受到严厉的惩罚或自残（如拔掉自己的牙或熊掌）。但是如果熊能够保持足够的理智而做出理性、冷静的选择，尽管不能解决问题——即无法确定对方的选择——那么沟通体系必须保证熊能够收到对方的讯息。对方也会向熊提出条件并保证其能够收到，因为只有这样，熊才能接受允诺（现在他看到了解决方法），对方才能给予确认。

博弈矩阵中的行为组合

假设博弈游戏中存在诸如威胁、承诺和允诺等潜在行为，那么，我们一定能够在前面提到的类似收益矩阵中，以战略选择的传统形式表示这些行为。但是前提条件是，这个收益矩阵必须能够包含各种行为选择方式。

首先，我们必须注意承诺、允诺和威胁通常具有一个共同特点：做出一个行为选择后，一方选择性地缩小了（明显或非明显）自己在收益矩阵中的收益范围。这也是行为的精髓所在。[1]

我们也可以说，一方提前公开其战略选择，作为对对方选择的回应。但是这不是简单的战略选择问题，一方必须对其战略选择负责。一旦失败，他必须承担相应的惩罚，即从自己的非相关收益中减去相应的数值。[2]

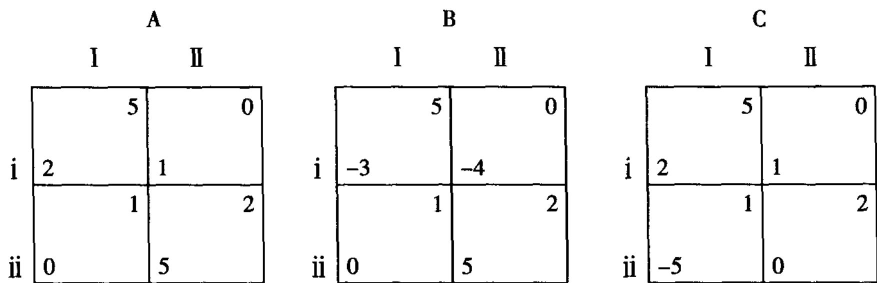
沟通是可以被拒绝的沟通。”他也谈到了“非法”窃听的问题。“一方可以从其他人那里了解有关情况。他知道他们是否知道这一事情以及他们什么时候揭露自己。”他还指出，即使窃听的人知道他们对其侮辱的言论是否进行还击则取决于他的窃听行为是否“合法”。

[1] 丹尼尔•埃尔斯博格的一些有关战略的著作中（第1章有所涉及）谈到了他对威胁或承诺形式的看法，即在收益矩阵中选择性地缩小自己的结果范围。

[2] 前提已经详细论述了威胁、允诺和无条件承诺。这里我们将讨论“互动功能”（reaction function）。假设行拥有足够的罚金用于自己的选择，他将留给列一个最大化问题，即列必须选择其第三战略，那么行将得到所有有利的结果。特别是，行赢得最有利的结果，即迫使列选择不低于最小的最大结果。这也是战术的规律化，双向或三项选择都可以被看做是承诺、威

特别是在图 5–7 的 A 中，行选择 ii 并从自己的结果中减去相应的数值 5，从而使 ii 成为其主要战略选择。这意味着，不论对方如何选择，行必须坚持自己的这一战略。结果如图 5–7 的 B 所示（如行选择 i 并接受惩罚 5，我们将得到图 5–7 的 C)。现在我们是否能够建立一个包含更多选择的矩阵，不仅包括前面提到的行和列的所有选择，如图 5–7 中的 A；也包括承诺、威胁和允诺战略等？毫无疑问，我们能够做到，只要我们规定双方的行为选择和双方选择的先后顺序。行有权提前做出自己的选择，而列拥有最初游戏的首先选择，即在行选择最终行之前选择自己的列。

在最初的游戏中，行是第二个做出选择的，现在拥有 4 个选择权，



A

	I	II
i	5 / 2	0 / 1
ii	1 / 0	2 / 5

B

	I	II
i	5 / –3	0 / –4
ii	1 / 0	2 / 5

C

	I	II
i	5 / 2	0 / 1
ii	1 / –5	2 / 0

图 5–7

胁、允诺或他们的组合（第 7 章中将对这一问题进行深入探讨）。

	i	ii	iii	iv	v
i	6 / 1	10 / 11	2 / 10	9 / 2	7 / 10
ii	9 / 8	4 / * / 12	0 / 25	1 / 20	15 / 3
iii	20 / 9	15 / 2	6 / * / 16	1 / * / 18	17 / 14
iv	2 / * / 6	10 / 8	7 / 7	4 / 5	3 / * / 20

即：任意选择 i ；任意选择 ii ;选 择 i 段的 I 和 ii 段的 II ;选 择 ii 段的 I 和 i 段的 II 。现在行可以做出自己的第一个选择，在选择的同时，他还可以采用上面提到的 4 个战略之一，即承诺 ii 时，任意选择 ii ;承 诺 ii 时，任意选择 i ;承 诺 ii 时，选择 i 段的 I 和 ii 段的 II ；承诺 ii 时，选择 ii 段的 I 和 i 段的 II ；行总共可以进行 2 个战略组合。

与行相比，列拥有最多 8 个战略组合：即第一次选择机会可能遇到三种情况。选择机会 I 和 II 时，三种情况分别为行承诺 i 、行承诺 ii、行没有承诺。

如果我们将这些战略矩阵化就会得到图 5-8。图 5-8 中 12×8 代表了双方在最初默式 （non cooperative ）博弈中的决策选择。例如，列的 8 个选择可以被看做是他给委托人的 8 个完全不同的指示。列如何选择取决于行是否或如何做出承诺。双方在最初游戏中的战术同样适应于新矩阵，因为新的矩阵完全按照同样的原理设计。双方的战略选择可称为

		i	ii	iii	iv	v	vi	vii	viii
		0- I 1- I 2- I	0- I 1- I 2- II	0- I 1- II 2- I	0- I 1- II 2- II	0- II 1- I 2- I	0- II 1- I 2- II	0- II 1- II 2- I	0- II 1- II 2- II
i	0; I – i ; II – i	2, 5	2, 5	2, 5	2, 5	1, 0	1, 0	1, 0	1, 0
ii	0; I – ii ; II – ii	0, 1	0, 1	0, 1	0, 1	5, 2	5, 2	5, 2	5, 2
iii	0; I – i ; II – ii	2, 5	2, 5	2, 5	2, 5	5, 2	5, 2	5, 2	5, 2
iv	0; I – ii ; II – i	0, 1	0, 1	0, 1	0, 1	1, 0	1, 0	1, 0	1, 0
v	1; I – i ; II – i	2, 5	2, 5	1, 0	1, 0	2, 5	2, 5	1, 0	1, 0
vi	1; I – ii ; II – ii	-5, 1	-5, 1	0, 2	0, 2	-5, 1	-5, 1	0, 2	0, 2
vii	1; I – i ; II – ii	2, 5	2, 5	0, 2	0, 2	2, 5	2, 5	0, 2	0, 2
viii	1; I – ii ; II – i	-5, 1	-5, 1	1, 0	1, 0	-5, 1	-5, 1	1, 0	1, 0
ix	2; I – i ; II – i	-3, 5	-4, 0	-3, 5	-4, 0	-3, 5	-4, 0	-3, 5	-4, 0
X	2; I – ii ; II – ii	0, 1	5, *2	0, 1	5, *2	0, 1	5, *2	0, 1	5, *2
XI	2; I – i ; II – ii	-3, 5	5, 2	-3, 5	5, 2	-3, 5	5, 2	-3, 5	5, 2
XII	2; I – ii ; II – i	0, 1	-4, 0	0, 1	-4, 0	0, 1	-4, 0	0, 1	-4, 0

图 5-8

"互动调整战略"（strategies of response and adaptation）。

如前面的游戏规则一样，现在我们将列在最初游戏中的选择分别标为 Ⅰ和 Ⅱ；行的选择标为 i 和 ii。此外，"2"表示行的承诺为行 ii；"1"表示行的承诺为行 i；"0"表示行没有做出承诺。在新游戏中，列的"战略"分别表示为 0–Ⅰ、1–Ⅱ、2–Ⅰ。它们的意义分别为：选择列 Ⅰ，但是没有做出承诺；选择列 Ⅱ，做出承诺行 1；选择列 Ⅰ，做出承诺行 2。对于行而言，每个战略包括 0、1、2 三个决策点，另外还有一对符号表示他针对列做出的选择。例如，1; Ⅰ – i ; Ⅱ – i 表示行的承诺为行 i，以后无论列如何选择，行必须选择行 i。知道图 5–7 中的 A 的结果后，双方也知道新图中的结果。现在，我们假设行和列分别指挥自己的委托人在新图中对抗（提供一个特定战略）。在发出指令之前，行和列分别对图 5–8 进行研究分析。实际上，他们进行的是默式博弈，他们的委托人作为信使。

这个新默式博弈的结局如何？我们能否确定这个博弈游戏的结局？如果能的话，那么结局是怎样掩藏于矩阵中的？对理性的选手而言，最初的游戏都有一个结局。（A）如果行承诺选择行 i 并接受违约惩罚 5，那么列可以看到无论自己怎么选择，对方都会选择行 i。假设列选择对自己最有利的左上格 i – Ⅰ，那么行选择行 i 将得到左上格的 2。（B）相反，如果行承诺选择行 ii（从其行 i 中减去 5），列就会选择 Ⅱ而非 Ⅰ，那么行将得到 5。（C）如果行没有做出承诺，列知道行将在其选择的列中选择最高分数。如果列选择 Ⅰ，行将选择 i，列将得到 5；如果列选择 Ⅱ，行选择 ii，列将得到 2；如果列倾向于 Ⅰ，那么行将得到 2。综上所述，行的最佳选择是行 ii。这是最明显的解决方法，得到的结果为 [5 2]，符合行的战略 2; Ⅰ – ii ; Ⅱ – ii 以及列的包含 2– Ⅱ的全部 4 个战略。（如果行第一个选择，那么列将一无所有。）这些是图 5–8 中的行 x 的关键格。（实际上，行的第一个选择涉及如何选择图 5–7 中的 A、B、C。）

那么，我们该如何给图 5–8 中代表解决方法的矩格或战略组合定性呢？它们代表一种可以被称作"最模糊意义的解决方法"（*a solution in*

the complete weak sense）的解决方法。[1]如在图 5–8 中，博弈双方可以通过放弃“主导”（dominated）行或战略来实现这一目标。如果每行的结果对选手行来说都至少与主导行的每一个结果同样有利或者至少与一个结果比较有利，那么，一行可以受到另一行的制约。根据这一原则，我们可以确定第一行受第三行的制约。（有人认为，行完全可以放弃第一行的战略，因为第三行在大多数情况下或至少在某些情况下比较有利。）第二个和第四个是同样的道理，以及除第十个以外的其余部分。第三个和第十个都不能控制其他行，我们暂时将其留下。与列相比，没有一列可以控制其他列，我们还可以留下第三行和第十行。（有人认为行可以二者都不选）选手列通过比较分析，最终选择第三列和第十列。现在很显然，第二列控制第一列、第三列、第五列和第七列。除去被剩下的行控制的列后，我们现在看看行 iii和行 x 。在最初的游戏中，这两个行谁也无法控制谁。剔除第三行后，我们剩下第十行，与四个列相交。四个交叉点的结果相同，这说明只要行选择第十行，列选择的四个战略就不合逻辑。（现在，行再次选择图 5–7 中 A 的第二行，这也是列所希望的。无论列怎么决策，这都会产生同样的效果。）[2]

[1] 请参阅卢斯和雷夫，pp.106~109。

[2] 我们淘汰列和行的顺序丝毫不会影响最终的“解决方式”。按照文中的程序，我们只留下第三行和第十行。随后，我们淘汰列 Ⅰ、Ⅲ、Ⅴ、Ⅶ 。由于第三行处于控制位置，所以我们淘汰第三行。留下的第十行与四个列同时交叉。此时，我们注意到仍有两列需要淘汰，即列 vi、Ⅷ，因为二者对列不利。我们应该淘汰第三行和列Ⅵ、Ⅷ，但是由于规则限制，我们只能先淘汰第三行，后淘汰列。从这个意义上可以说，我们解决方式取决于规则的强制性。不管我们得到两个相同的结果，还是四个相同的结果，都依赖于强制性的规则。在这一阶段，列明白自己没有必要理论，因为行的选择非常明确，即，无论列怎么选择都将得到一个不合逻辑的结果。当列发现这一点时，他会留下哪个列一定程度上取决于他在决策过程中的选择空间。（如果列的选择存在沟通成本，那么列可能仅选择战略 2–Ⅱ，从而存在多种选择应付行的战略 0 或 1。假设情况相反，如果行的战略被错误地记录或沟通，或者非明智地选择，那么列只有选择 0–Ⅰ 才能化解风险。在后一种情况下，他考虑了第三行因素，尽管存在第十行。我们进一步假设，他怀疑裁判倾向于行选择行 Ⅴ，尽管行选择了其他行。列可以进一步将自己的选择缩小到 0–Ⅰ、1–Ⅰ、2–Ⅱ，最终的解决方法就是第十行和列 Ⅱ的交叉点。因为这是最好的选择。一般而言，由于各种错误的可能性和不同的战略选择成本也不同，所以会出现很多问题，并可能得到不同的结论。第 7 章和第 9 章探讨的问题涉及许多随机行为、错误或信息误差因素。）

这也是最初的连续行为博弈（sequential move game）的解决方法在静态博弈（“无行为”或同时默式博弈）中的体现。这一解决方法要旨在于，主导因素必须反映每一阶段选择的战略和淘汰被控制的次要战略。当连续行为博弈具有明确结果时，这看起来更像是发展后的默式博弈的普遍解决方法。筛选行和列可以被看做是一个逆向思考过程。首先，思考做出理性的最后选择来确定一个宏观方向。其次，考虑最后一个选择是以前一个选择为基础。最后，思考倒数第二个最佳选择，从而做出博弈的第一个最佳选择。

在欣赏发展后的抽象“超级博弈”（supergame，一般形式）如何将威胁、承诺和允诺融合贯通的同时，我们必须强调我们仅仅通过研究一般形式的博弈还无法完全认识到这些策略。我们的研究对象是以沟通和强制体系为基础的策略和行为选择的时间，但是这些因素在一般博弈场景中都不存在。我们需要建构的理论必须能够将对行为结构博弈一般要素的研究系统化，而一个过分抽象的博弈模式不具备这些因素。[1]

表示连续行为博弈的矩阵不仅仅有助于强调博弈的标准定向性，因为即使是默式行为也不能削弱战略博弈的普遍性。一次威胁行动之所以成功并决定结果，其惟一原因是威胁迫使对方做出对威胁方有利的选择。对方可以保留自己最初的选择权，而这种选择权很大程度上依赖于威胁方的最终选择；而威胁方的第一选择——即是否做出威胁承诺——取决于他对对方对自己行为的预期判断，这样就保留了博弈本身具有的预期互动特征。和无条件承诺或众多行为选择存在时的广义上的“互动功能”一样，威胁只有当威胁方调整自己的动机结构以影响对方的预期判断时才能发挥效果。

[1] 以特定的选择矩阵来表示某个博弈场景通常不是一个可行的分析方法。众多行和列已经成为一个天文数字，既使简单的博弈场景。为了更好地说明问题，我们现在考虑一个 3×3 矩阵。列首先选择，行有权无条件地选择一个部分或完全具体的战略。最后，为了研究如何化解威胁，列仍然有权首先选择，即列有权无条件地首先选择，随后，行有条件地随意做出选择。最终，列选择一个列后，行选择一个行。为了保持游戏的简单性，我们不必限制惩罚的力度或引入不确定因素或非完全沟通体系。这样一个在扩展形式下并不难于分析的简单博弈，却会出现无限多的列（10的 100次方）。

矛盾的战略优势

当然，根据推论原则，假设收益矩阵一开始就告诉双方选手的一方，其为取胜所要付出的分数，那么，他不费吹灰之力即可取胜。（第2章最后一段论述了这一观点，即谈判中的劣势也可转化为优势。）显然，一个博弈论原则不可能将整个混合博弈概括为：对一方而言，部分或全部潜在不利结果的恶化可能转化为一种优势。这也解释了为什么严厉惩罚答应黑邮件的人是对其人身安全的保护；为什么面对强敌，背水一战往往能转败为胜；以及为什么年轻女孩总为寻找心中的白马王子而拒绝别人。[1]

据非官方消息报道，朝鲜战争期间，财政部冻结中国在美国的资产的同时，也冻结了一些非中国的资产，并将其作为防止美国华人的中国家属受到不必要威胁的手段。对旅居美国的华人而言，禁止他们向中国转移资金增强了他们抵制来自中国的威胁。人为地增加资产非法转移的困难度、建议严厉惩罚非法转移个人资金，或冻结那些所谓的社会主义同情者的资产等措施有效地保护了潜在的受害人，预防了强式威胁的发生。

[1]这也解释了为什么有时好心没好报。一个可以使一方安全地做出某个选择的允诺，让我们相信他一定会这么做。这样，我们就会以这个允诺为基础做出对其不利的选择。如果可能的话，我们可以将价值加到别人的结果中，那么这将对其十分不利。在下面的矩阵中，假设行首先选择，他可以赢得列和 7分。如果行向列单方面承诺，一旦出现 i－II，他将从自己的分数中减少相应分数补偿列。如果他承诺补偿 2分，他将得到 8，列得到 3。否则，行不能选择 i，结果将是 ii－I，双方分别得到 1和 10。列当然希望行还要承诺。（如果勒索者也无法令他降低自己的要求，即使加上勒索的罚金，也小于他威胁后受到的损失。他也许会为被威胁方支付罚金。因此，这就保证了被威胁方对威胁的反应。这样，威胁成功，而且对被威胁方不利。）

0 \ 2	10 \ 1
1 \ 10	2 \ 0

转化为

0 \ 2	8 \ 3
1 \ 10	2 \ 0

《旧金山对日和约》中的第 26 条也体现了这一原则。该条款规定，日本在向第三国做出领土让步之前，必须征得美国政府的同意。 1956 年，当有报道称日本人迫于前苏联人的压力将做出领土让步时，美国国务卿杜勒斯在新闻发布会上援引该条，并表示他已经“提醒日本政府遵守这一条款”。[1]很显然，杜勒斯这是通过新闻发布会向前苏联人发出了警告。杜勒斯还向日本人说了一句大家都比较熟悉的谈判辞令“如果我为你这么做了，我必须以同样方式对待别人”。在前面我们也提到了这一谈判辞令。任何欺骗美国的行为都将受到严厉惩罚，这有力地保证了该“承诺”的有效性。（矛盾之处在于，美国不会让日本人得到利益，除非这一策略失败时，美国愿意发挥这一条款的作用。）[2]

“战略行为”

虽然博弈双方之间的互动决策和行为互动是战略博弈的精华所在，但是我们还是有必要给“战略行为”定性。战略行为是指，博弈一方以对自己有利的方式影响对方对自我行为的预期判断，从而达到影响对方决策的行为模式。战略行为意味着一方在限制对方行为之前必须首先限制自己的行为。其目标是为自我建立一个行为模式，并将这个行为模式推销给对方（包括对对方的有条件反应）。这一行为模式给对方留下一个最大化问题，并迫使对方做出对自我有利的选择，同时也破坏了对方采取同样行为的能力。

混合博弈与完全冲突（零和）博弈相比，惟一的区别在于前者强调博弈选手制定战略并将其传达给对方的重要性。而零和博弈的要点在于博弈双方不能进行沟通以及一方的决策不能被另一方察觉。[3]只有混合博弈强调战略行为的重要性，并建立双方互动行为模式的优势。

[1] 美国国务卿杜勒斯在新闻发布会上的讲话。*The New York Times*（August 29，1956），p 4。

[2] 一方恶意提案通过的反安乐死法将削弱对方的立场。“对那些病人膏肓、想入非非的老人来说，安乐死无疑是一个了此一生的好方式。”（参见 Jorn Beavan，“The Patient’s Right of Live and Die”，*The New York Times Magazine*，August 9，1959，pp.14，21-22）。

[3] 范·纽曼和摩根斯特恩说（p.147）：“我们将一方能够破解对方的战略这一危险现象作为考虑的重点。”

当然，在零和博弈中，博弈一方如能促使对方相信某种博弈模式的存在，这对其自身也是一个有利条件。在混合博弈中，一方如能按照规则约束自己的行为，并能预测自己会成功时，他就非常乐于将这一信息传达给对方。

在混合博弈中还存在一个看似矛盾的现象：一方的装疯卖傻、故意回避如能赢得对方的承认和信任，那么，这对他来说也是一个优势。这个看似矛盾的现象通常出现于合作或化解威胁的过程中，而不会出现在零和博弈中。同样，在零和博弈中，由于理性的博弈双方掌握完全资迅，即使拥有第一选择权的一方也并非一定具有优势［范·纽曼（Von Neumann）和摩根斯特恩（Morgenstern）的“少数人的博弈”］。但是在混合博弈中，情况恰恰相反。

第6章

博弈论与实验研究

前几章讨论了博弈论的研究方法问题，并得到了几个初步结论。结论之一，结果导向的数学结构分析方法不应成为博弈论的主导研究方法。结论之二，研究过程中，我们不应当将问题过于抽象化。如果我们改变博弈场景具体变量的数量，那么，我们有可能改变了博弈的特性。因为其中有些变量可能具有重要价值，如博弈双方对彼此价值观的无知等。在一般情况下，这些具体语境中的变量能够引导博弈双方实现明确结果，至少是双赢的结果。在前面的案例中，霍尔摩斯（Holmes）和莫日娅蒂（Mariarty）之所以能够在同一个车站下车，很大程度上取决于博弈场景中的变量，而非问题结构本身。这些变量或者存在于火车上或者存在于火车站，或者双方拥有共同的背景，或者当火车停车时，他们从高音喇叭中得到某种启示。尽管我们无法以科学方式得出，究竟是什么因素导致双方合作，但是我们必须承认一定存在某种因素决定双方的结果，或许是抽象研究中认为的非相关因素。结论之三，当沟通方式具有某种优势，博弈双方对彼此价值观或战略选择缺乏了解，特别是博弈结果依赖于博弈双方的一系列行为和举动时，经验因素往往成为混合博弈研究的关键因素，也是最适应的方法。但是，这并非意味着人们如何进行混合博弈只是一个简单的经验问题，特别对需要高智商的复杂博弈场景而言。因此，我们可以肯定：博弈的成功原则、战略原则和标准理

论的命题并非经验分析方法所能得到。

在零和博弈中，研究人员往往只分析博弈双方中一方的理性因素和决策选择。事实上，博弈双方都能做出理性选择，但是最小最大策略将这一场景变为一个博弈双方必须单边决策的过程。双方之间不需要任何形式的共识、思想撞击、任何暗示、任何直觉或互谅。总之，在零和博弈中，博弈双方不需要任何社会性认知。然而，混合博弈不仅需要双方的互动，还需要多方的互动。双方必须就相关问题进行沟通，至少取得某种共识。通常情况下，双方之间需要某种行为活动，无论是潜意识的还是默认的。博弈双方能否取得满意结果取决于双方之间的社会认知和互动程度。甚至是两个完全隔离、无法进行言语沟通，甚至不知道彼此姓名和身份的选手也一定需要进行心理沟通。

结果，研究人员无法通过内省、自明再造整个决策过程；更无法建立一个双方或多方决策单位互动的决策模型。在这个模型中，决策单位的预期和行为完全来自于标准的逻辑推理，假设研究人员知道某个理性决策单位进行决策的标准，他们也能逻辑推理出其决策结果。然而，研究人员无法通过单纯的标准分析推断出双方进行沟通的内容，因为这需要两个人的努力。（两个研究人员或许能做到这一点，但是前提是双方必须是实验客体。）毕竟理解一个暗示与解读一个标准的沟通行为或解决一个数学问题有着本质的不同。前者涉及某个特定语境中的信息，一方决策单位希望将该信息传达给另一方决策单位，并使其产生某种印象或联想。如果没有相关经验的话，一个人很难推断出某些非零和博弈行为的真实含义；而一个丰富经验者则完全可以通过逻辑推理，理解某个特定语境中的信息，如一个笑话或幽默。

为了说明这一点，我们向大家举一个例子。假设两个人看到同一片墨迹，如果他们都试图知道对方对墨迹的理解，他们能否产生同样的印象或联想呢？问题的答案只能通过实验才能知道。但是，如果他们做到了，双方就可以实现完全标准博弈理论所不能做到的事情；而且，他们做得要比后者想像中的更好。如果情况真是如此——他们能够超越完全标准的博弈理论的范畴——那么，甚至是标准的描述性战略论也无法建立在完全标准的逻辑分析之上。这样的话，我们就不能通过假设某种思

维过程的存在来建立一个描述性理论或说明性理论。因为这一思维过程是理性选手所无法实现的，并且属于“理解暗示”类型的思维过程。博弈选手个体或集体在实际中的行为能否超越完全标准博弈理论的预期，从而否定理论战略原则的作用，这是一个需要验证的问题。[1]

在此，我们必须再次强调，在零和博弈中之所以没有出现这一现象，很大程度上是因为博弈双方没有同时有效利用社会性互动行为或至少一方故意破坏所有的社会沟通。而在非零和博弈中，由于博弈结果受到多种因素影响并要求双方必须合作互谅，所以，一个选手根本无法故步自封，完全脱离社会环境。当然，如果完全静音的收音机也无法促进双方进行有效合作，那他就不能随意地丢掉助听器以堵视听；假设信件已经寄出，那他就不能故意回避，因为对方期望他打开信件并采取相应的行为。

这时，我们不禁要问，难道博弈论的触角已经扩散到整个社会心理领域了吗？难道博弈论将我们引入到了一个更狭隘的博弈论领域？是否存在混合博弈需要实现合作的普遍规律，而且能够被实验观察所发现，并对谈判具有普遍的指导价值？尽管我们还不能肯定这一点，但是我们

[1] M.M.弗勒德 （M.M Flood）成功进行了一个有关博弈战略沟通认知问题的实验。实验中，弗勒德设计了一个 2×2 非零和博弈矩阵模型，双方选手拥有 100 轮连续的默式选择机会。矩阵的特别之处在于，选手双方每次只能选择某个特定的格才有可能获胜。但是，如果合理分配胜利机会，那么，双方必须合作处理某 2个或更多的特定的格。双方实现合作的惟一方式是在游戏进程过程中，双方不断协调自己的选择。“沟通”阶段——或者在以后的阶段中，犯规一方将受到报复模式的惩罚——需要双方必须为此付出一定的代价，因为双方失去了收益的非协作机会。参见 M.M Flodd，“Some Experimental games,” *Management Science*，5：5~26（October，1958）。

双方之间如何进行有效沟通或释义对方行动模式中暗示，很大程度上取决于双方之间对共同行为模式的共识——共同建构某种模式的能力——并非前面提到的完形心理学家所进行的实验过程。尽管完全标准的沟通理论可能推导出理性选手应该得到的最低的有效沟通标准，但是双方选手能否得到更好则是经验问题。双方如何理解暗示和什么样的暗示容易被理解完全是社会认知问题，或许必须经过实验研究。（假设拍卖会的两个竞拍人意识到相互竞价，双方都将受到损失，因而努力就某种模式达成共识，从而避免自相残杀，将得到的利益和机会均等瓜分。）

相信一定存在某个值得研究的领域。即使我们不能找到一些普遍规律，至少我们可以推动现有研究成果的发展。至今为止，实验领域的博弈论研究还是取得了巨大成果。

想想前面提到了的地图游戏和象棋游戏，他们都属于非零和博弈。类似的还有“有限战争”案例。博弈双方取得成功的原因，很大程度上是因为他们都竭力避免自相残杀的战略。这里有一个案例再次证明，博弈双方能否避免两败俱伤取决于二者的心理默契。这个游戏包含的一些细节在游戏一开始就映入了参与者的脑海，如地图或版图分布、各种因素的名称、游戏本身具有的传统和先例以及精美寓意的背景。如此复杂的游戏要求双方必须寻求某种社会认知，并进行有效的沟通。假设我们已经克服了建构该游戏的技术问题，那么，我们将进一步思考游戏需要回答的问题和验证的假想。

或许有人会问，当出现下面的情况时，是否能够找到一个成功解决问题的方法，即双赢的结果？ （A）双方能够进行完全有效的沟通。(B) 双方只能进行行为沟通，不能进行言语沟通。 （C）双方之间的沟通具有不对称性，一方发出的资讯量超过另一方的接受量。尽管我们不能保证一定存在一个放之四海而皆准的答案，但是我们还是有机会发现一些有关沟通作用的普遍规律。目前学术界对博弈双方的有效沟通、单方面声明、双方之间的默式沟通能否增加有限战争的可能性存在争论，这一切都充分说明了这一问题的重大意义。[1]

另一组有关有限战争或国际战争的问题是，双方在什么场景中更有可能实现明确、有效的结果，是双方将游戏中的各种隐含条件——各种行为或因素的名称——熟记于心时，还是当这些新奇、陌生的条件令双方无法达成共识时？理性对手能否在东南亚进行一场有限战争——广义形式的博弈——无论是常规武器还是核武器，或者在月球上与一个未知

[1] 为了避免任何可能的误解：作者并不认为实验可以完全模拟战争状态，或者有关有限战争的实验结果可以直接应用于现实世界。通常这类试验属于“基础研究” （Basic research），主要涉及问题的认知和沟通方面，而非动机——除非动机影响到了社会认知。实验结果可能具有实际意义，被直接应用于现实。而且，随着现在的理论研究通常在默式实验博弈基础上研究有限战争中的沟通作用和有限条件的类型，这一可能性日益增大。

敌人进行有限的生化战争？这些都是非常重要的问题，因为他们都涉及博弈论的核心问题。没有充分的实证，谁也无法回答这些问题。当然，我们有理由相信，博弈双方具有足够的智慧来忽略这些博弈场景的细节。这些细节的意义在于其有助于博弈双方进行有效的沟通，实现互谅。

与缺乏共同文化背景和个性的博弈选手相比，具有共同文化背景和相似个性的选手是否更容易达成共识，实现合作呢？假设博弈双方是经验丰富的老手，或是两个新手，或一个老手和一个新手，那么哪个组合更容易找到解决问题的方法？换言之，在博弈中，老手是否比新手更具有优势？

在类似的博弈场景中，开局重不重要？如果博弈双方在初期没有找到游戏规则，以后他们是否还能找到？假设博弈双方都受到规则或“有限”武器和资源的限制，除非情形需要可以适当放宽；或者双方为了避免以后出现规则愈演愈松的局面，在游戏开始就实施宽松的条件，那么他们在哪一种情况下更容易实现互赢？

调停人在这类博弈中将产生什么影响？什么调解方式更有效？博弈结果与调停人存在利害关系对双方而言，是利大于弊，还是弊大于利？如果调停人不公平，那么不公平到什么程度不会影响明确、有效的结果？

如果在这样的博弈场景中，选手双方能够在某些问题上为彼此记分，那一定非常有意思。例如，谁更加努力进取或具有合作精神？双方认为什么样的规则有效？谁是赢家（由于双方对彼此价值观缺乏了解，只能进行某种释义）？什么时候是博弈的转折点，或必须变化策略，或对方的某个行为将被释义为报复行为或新条件？

由于“报复规则”从本质上讲具有诡辩性；由于阻碍有限战争的因素主要来自双方的心理因素和诸如传统等社会因素；由于诡辩和传统的客体通常不能控制博弈场景，在美苏进行核报复的时候，欧洲已经实现了有限核战争；或在没有种族冲突的地区语言学校发生爆炸事件；或在特殊的行业引入非价格竞争机制。所以博弈论的经验部分需要穆扎夫·舍里弗（Muzafer Shelif）的实验研究。他发现，如果实验不存在判断标

准，那么，博弈双方会自己制定相关标准；如果存在这样的标准，那么，一方总是设法发展这一标准并影响对方。双方之间存在一个价值互动的过程。一方在形成自己价值判断标准的同时，也试图协调自己与对方的价值观。当博弈场景本身提供的“客观”标准无法形成一系列规则时，即博弈本身存在不确定状态时，双方必须自己制定有关的标准。只有这样，双方的行为互动才能“合法化”。[1]双方可以通过无意识的合作就某些条件、挑战性和肯定的行为或合作性姿态达成共识。他们还必须共同制定对违反规则的行为的惩罚标准。[2]

例如，某个博弈场景可能将双方中的一方定位为侵犯者；或者为领土的重新划分提供某些背景资料；或者在版图的某个部分赋予其中一方某种道德权力。当然，这些背景资料不会影响博弈场景的逻辑和数理结构，因为它们只具有提示参考作用。我们也可以建立一个与前面游戏相同的版面，并验证如果双方知道前面的形势结果后，现在的结果是否会发生变化。如果选手希望以最初的游戏格局来制定标准，那么，一个背景故事的出现可以扭曲这些标准，[3]因为前者促进双方从其他角度假设条件。

[1] 实际上，有关博弈双方制定标准的精典案例是 1957 年的裁军谈判——表明了这一过程的研究价值，有关最后各方一致同意核查地区必须是北极的某个饼形区域。

[2] 有人希望博弈理论家能够区分与博弈论相关的心理实验和其他社会心理学的差别；也希望存在一个战略论，而非全部是冲突行为。但是，谁也无法肯定二者之间一定存在界线。例如，非博弈论意义上的“敌意”或许仅是感情或性格特征。但是，如果是在博弈场景中，敌意可能会成为一方理解另一方的意图的障碍，那么敌意就成为“沟通结构”的一部分。杜切（Deutsch）做了一个相关实验。他让两名选手连续两次进行默式非零和博弈（以矩阵形式），双方可能进行合作选择或非合作选择。选择不合作的一方比选择合作的一方多一个机会，第二轮中可以选择是否合作。但是，“如果一方的预期判断没有被对方认同，他们可以认为双方的选择出现分歧，并就游戏的规则缺乏共识……在这一组，一方对另一方选择的承认意味着加强了对方此前的消极意图。”参见 Morton Deutsch，*Conditions Affecting Cooperation*，Research Center for Human Relations，New York University，1957.（*The Journal of Conflict Resolution* 刊登了一篇以本文为基础的题为“Trust and Suspicion，”的文章，并不包含这里引用的观点。 2：265~279［December 1958］）。

[3] 第 3 章提到的所得税问题说明了建议权的重要性（pp.62~65）。

更有趣的是，一方能否发现对方正在考验他的意志，或对其实施激将策略呢？只有当博弈过程中存在某个具有提示意义的点时，我们或许可以发现这一过程，如双方都意识到各自在博弈过程中的某个点为自己建构了一个角色和尊严。

这一博弈场景的另一个值得研究的地方是，双方行为或价值观都涉及渐进主义（incrementalism）意义。象棋游戏和两军对峙等博弈场景也属于类似情况。以象棋为例，如果棋手双方轮番走子，每次只能走一步，那么整个游戏在渐进过程中完成。双方以小积大，经过不断地走子，最终改变整个棋局的走向。双方的每一次失误都会削弱自己的有利形势，最终导致棋局的失败。如果双方能够进行言语的沟通，那么二者一定会唇枪舌剑、争论不休，并避免两败俱伤的招数。现在，假设双方每次可以走数子，任意方向，任意距离，而且规则规定结果必须对双方或其中一方不利，那么，游戏不再是循序渐进的，而是激进式的，甚至会出现突袭。棋手双方也只能看到眼前的局势，而无法预测下一步或下两步。双方之间也几乎没有暂时妥协的机会，或互信传统，或一方主导另一方妥协的局面，因为游戏的进程远远超过了双方的预测和理解。但是，一个更加渐进的游戏是否更好地促进双方的合作，或增加了游戏的风险系数？这是否取决于选手双方的自身特征和我们在游戏中设置的暗示条件？选手双方的行为或价值观念的渐进主义是否发挥主导作用？这两个因素能否相互补充，从而在一个游戏中引入渐进主义能够弥补另一个游戏缺乏带来的影响。有关的争议还很多，诸如核武器在有限战争中的作用、在威慑存在，双方不愿进行现代战争并希望战争地区化时，进行突袭将有什么意义，以及西欧大陆能否进行一场有限战争。一旦实验观察确立了实证的标准，我们就可以有效地分析渐进主义。[1]

这些博弈场景大多涉及两个选手，除了某些场景涉及调解人之外。两人以上的独立选手也可以进行类似的游戏、每人都为自己的收益负

[1]“有限战争不仅必须有寻求阻止最极端的暴力行为的方式，而且必须寻求阻止延缓现代战争的爆发以促进政治目标与军事目标的联系。一旦这一关系消失，任何战争都有可能在不知不觉中演变为一场全面战争。”（参见 Henry A.Kissinger，*Nuclear Weapons and Foreign Policy*，New York，1957）。

责。作者推测——至少在成功选手之间——多人参加更有利于经验结论的涌现。简而言之，像暴乱和联盟这样的现象都需要进行经验研究。与博弈论研究通常使用的对称性等方法不同的是，人为地将非对称性、先例、行为次序、非完全沟通结构以及各种隐含因素引入博弈论将有助于博弈理论的发展。各种非对称性因素和非完全沟通体系对联盟形成的影响将其引入系统的实验研究。[1]

[1] 亚历克斯·巴弗拉斯曾经进行了一个完全合作博弈实验。实验中五个选手分别传递手中的棋子，直到形成五个不同的方框。问题在于一个选手摆好一个方框后，剩余的棋子不能满足剩下的四个方框。亚历克斯·巴弗拉斯感兴趣的是，当成功可能实现时会发生什么事情。“当一个人摆好一个方框后，很显然他不愿意将其拆掉。他是否脱离的终极目标很大程度上取决于他是否准确地分析了整体形势。鉴于此，沟通模式发挥了很大作用。最初的进展表明，重拆和坚持的斗争非常激烈。如果没有足够的沟通，实现最终目标将非常困难。”参见 Alex BAvelAs, *Communication Patterns of Task-oriented Groups*, D.Cartwright 以及 A.F.Zander, *Group Dynamics* [Evanston, 1953], p.493)。其他相关实验成果还有 Charles E.Osgood, The Biased Perception of What Is Equable, Suggestions for Winning the Real War with Communism', *Journal of Conflict Resolution*、3: 304~305 (December, 1959)。

第三部分

战略：充满随机性

第 7 章　允诺和威慑的随机性
第 8 章　产生随机因素的威慑

第7章

允诺和威慑的随机性

完全冲突博弈论（零和博弈）中，随机化战略发挥着核心作用。可以毫不夸张地说，在过去的 10~15 年里，随机行为因素一直是博弈论研究的核心问题。[1] 在二人博弈中，随机化的本质是为了避免对方掌握自己的行为规律，防止对方通过分析自己的行为踪迹而掌握自己的行为规律，最终达到迷惑对方的目的。然而，在冲突和共同利益并存的博弈场景中，随机化并没有如人们期望的那样发挥如此重要的作用，而是完全不同的作用。[2]

[1] 范·纽曼在谈及“最佳战略基本定理”——该定理认为，所有包含有限数量战略选择的零和博弈都能够使最小最大与最大最小产生均衡（“解”），如果混合战略允许的话——时说：“据我所知，所有博弈理论都是建立在这一定理基础之上的……我认为，除非‘最小最大定理’得到验证，否则任何研究成果都毫无意义。”（“Communication on the Borel Notes”, Econometrica, 21: 124~125 [JAnuary 1953]）。

[2] 然而，有人可能将零和博弈中的混合战略看做是将战略持续性引入非连续博弈的手段，从而赋予后者重要的完全战略鞍点。但是，零和博弈中的混合战略与非零和博弈中的混合战略并没有区别。一方可能通过掷币来阻止对方准确地猜测是正面还是背面。一方也可能采取中庸战略，通过掷币来平衡正面与背面的概率。两种方式都有效。如果第二种方式过于复杂，那么一方可以更好地切中问题的本质并展现给选手。第一种方式提醒我们，这一问题（尽管引入随机化）仍然能够阻止对方预期判断出我们的实际战略选择。决策机制、决策的记录和沟通程序，以及随机过程的结果要求的任何前期准备必须避免对方情报人员的接触。

在那些博弈场景（非零和博弈）中，随机化的主要任务并非防止对方预测自己的战略。正如前面所提到的，在那些博弈场景中，一方不是想尽办法掩蔽自己的战略，而是努力使对方准确无误地预期判断到自己的战略。

当然，在一些规模更大的博弈场景中也可能存在零和博弈选手。尽管有限战争中，一方或许关心能否与对方进行有效的沟通，而非掩蔽自己希望对方关注的有限条件，但是在那些有限条件中，一方也可能通过随机方式为自己的飞机诡辩，从而最大限度地迷惑对方。[1]双方可以进行有限信息的沟通，并在此基础上达成有限协议，但是双方谁也不能将信息完全告诉对方。以军控协议为例，军控条约必须在双方有限技术手段监督下实施，才能向双方展示彼此的军事力量是否遵守了条约，从而避免一方采取突袭行动。

但是，随机化在传统的零和博弈中却发挥着截然不同的作用。随机化的主要任务就是分配不可分配的对象和促使不相称的对象和睦相处。当随机化实现这一目标时，往往采用抽签方式决定双方应得的“预期价值”（expected values）。我们可以掷币决定谁将拥有分配的对象；如果这也做不到的话，那就只能以“双赢或零”来决定。当我们希望从一批合格者中选拔优秀人才进行长期服兵役时，我们可以通过抽签公平地分配公民义务，选择被征召的入伍者。

发挥这种作用的随机化显然与允诺存在密切联系。如果需要允诺的价值供大于求或不可分配时，那么抽签方式则有可能使允诺的预期价值大大缩水，当然也可以降低允诺主体的承诺成本。对一个人而言，危难之时得到的帮助允诺与一点点实际帮助一样具有重要的意义。（如果提

[1] 在某些特殊情况下，双方之间通常面临保密与沟通的两难境地。如果双方为了表明自己实施威胁的决心和能力，那么他必须以事实向对方证明。当然这些事实也可能向对方传递有利于其抵制威胁的信息。例如，假设我们希望向对方证明自己的新式武器超越对方的防御能力，那么我们必须向对方展示新式武器或其一部分，或者提供技术资料证明武器的性能。这样虽然对对方产生了威慑，但同时也有助于对方提高自己的防御能力。假设我们希望在某个地区展开一场地区性战争，那么我们必须提前将部队派往该处，这无疑为对方提供了战斗的具体地点，从而避免对方大海捞针，四面出击。

供的帮助正是对方危难时所必需的，那么雪中送炭的效果将更佳。）

但是，允诺在这一方面与威胁不同。不同之处在于，前者成功时需要付出高昂代价，而后者失败时才需要付出代价。一个成功的威胁就是不需要付诸实施的威胁；与此相反，作为一种诱导，如果我允诺的部分超过我必须付出的，那么允诺就是成功的，因为我的实际付出远远超过必须做的部分。但是，威胁的代价看起来不是有些高昂，而完全可以用奢侈来形容。如果我威胁殴打并折磨我们，你也许会继续顺从，因为我既不会折磨我们也不会杀死我们，这个错误没有造成任何损失。假设我现在只有一个足以将我们大家炸死的手榴弹，但是我只想使用催泪弹，那么如果你还不顺从，我只能被迫将威胁使用手榴弹降到威胁使用催泪弹，因为威胁手榴弹可能走火将我们一起炸死。但是在允诺情景中，这样的必要性并非如此明显，因为允诺的价值也可能丧失殆尽。

如果做出威胁承诺与实施威胁的成本相同或大规模威胁的成本远远大于小规模威胁，那么，威胁的程度就成为一个必须考虑的因素。如果威胁使用催泪弹已经足以达到目标，那我也就没有必要再威胁使用手榴弹；如果使用催泪弹比使用手榴弹成本更低，或者我必须展示手榴弹以增强威胁的效果，那么，最好还是选择使用催泪弹。但是，当手榴弹成本更低时，那就另当别论。

有趣的是，大多数威胁的最大成本是不得不实施威胁，普通意义上的“成本”则往往并不是决定性因素。

失败的风险

失败的风险迫使威胁实施者往往选择比较保守的威胁方式。假设必须以恐怖方式进行威胁，那么威胁方更倾向于采用抽签方式——仅仅威胁可能采取某种行为，除非对方接受自己的条件，而不是承诺一定会采取某种惩罚行为。

为了方便理解，这里我们引入一个矩阵图 7-1。在图 7-1 中，列首先选择，行有权对列实施威胁从而达到影响其选择的目的（X，Y 均为正数）。一种情况是，如果列选择列 II，行将以行 ii 威胁列。如果行没有威胁列，那么，列知道自己选择 II，行就会选择 ii。如果行对列实施

	I	II
i	1, 0	0, 1
ii	0, 0	−y, −x

图 7-1

了威胁——假设行做出了威胁承诺，列知道这一威胁的存在——那么，列选择 II 对双方都没有好处，则列更倾向于选择列 I。

现在的情况是，行认为绝不会发生对自己不利的情况，即使是自己完全误解了列的意图。正如图中所示，列的选择也许会与大多数人的选择相同。然而，拥有完全不同偏好的少数人可能更倾向于右下格而非左上格。同样，行可能做出了威胁承诺，但是却没有将这一信息准确地传递给列，并使其相信在这种情况下，列将会错误地忽视这一威胁的存在，做出错误的选择，从而导致对双方都不利的结果出现。同样，列也可能按照自己的承诺选择 II ，但是却没有及时准确地将这一信息传递给行，那么，行就不可能知道列选择了 II 。在这种情况下，行的承诺只能造成两败俱伤的结果。总而言之，无论何种情况，威胁都有失败的可能。如果我们将这一因素考虑在内，那么行就有充分理由相信，右下格的收益或许并非那么不利。

如果行将自己局限于“完全”战略——假设行能够准确无误地做出威胁承诺——那么，行只能期望右下格中的结果具有积极意义。但是，如果行将其威胁随机化，那么，他就能将自己的威胁成本降到某个具体水平。例如，假设行威胁列，如果后者选择 II ，前者就不一定选择 ii，而有可能选择 i 。这样，他就可能以更小的威胁成本实现自己的目的，即迫使列选择 I。

我们还可以进一步向大家具体解释。 P 代表威胁挫败的可能性，不管出于什么原因。（为了易于操作，我们将 P 设为自变量，并且不受行的战略选择的影响。）π 表示行的威胁选择 ii ，如果列选择 II 的话。换而言之，如果列不服从行的威胁，那么行的选择将对双方都不利。所

以，为了实现双赢，列只能选择 i。那么，行选择的 л 的值究竟应该为多大最合适？

首先，π 的值大到什么程度，威胁方才能取得完全成功（排除 P 因素的影响）？现在面对 л，列遇到了一个难题：如果列选择 I，那他将得到 0；如果列选择 II，那么结果可能是 1 和 –X 的平均值，两个砝码分别为（1–л）和 л，这对其非常有利。如果 1 和 –X 的平均值小于 0，那么列可能选择 I（受 P 的影响）。尽管如此，列还是会选择 II。威胁的成立条件是：[1]

$$0>(1-л)-лX$$

$$л>\frac{1}{1+X}$$

其次，我们假设由前面公式得到的威胁可能受（1–P）的影响而成功，也可能受 P 的影响而失败。如果威胁成功，行的收益将是 +1；反之，如果失败，其收益则为 0 和 –Y 的平均值，两个砝码分别为（1–л）和 л。当威胁条件充分有效时，结果为：

$$(1-P)+P(0-лY)=1-P-PлY$$

显然，这个值过大，而 л 相对过小。因此，为了满足第一个条件，行应该选择 л。要使威胁值得一试——即行不实施威胁时所期望得到的收益，期望值大于零——那么 л 的值应为：

$$1-P-PлY>0$$

$$或 \quad \frac{1-P}{P}\cdot\frac{1}{Y}>л$$

[1] 由于上面的分析仅仅建立在双方绝对结果之差的基础之上，因此没有强制双方将自己的优先选择为+1，第二个优先选择为 0。1/(1+X) 的完全理解应该是：（1）与（2）的对比。（1）为列的右上格结果与左上格结果之差；（2）为（A）列的右上格结果与左上格结果之差与（B）列的右下格结果与左上格结果之差的总和。公式的简化反映了采取尺度表示法的优势，因为仅仅一个变量足以表示 3 个值之间的关系。（后面的问题将涉及左下格，而且 4 个结果相互联系，因此需要两个变量。但是，如果左下格结果等同于其他中的一个，那么仍然可以说明问题。我们将得到非完全信息，更多的结果是 0 和 1。）详见 A.A.Alchian，“The Meaning of Utility Measurement,” American Economic Review，43：26~50（March，1953)，或者卢斯和雷夫的案例，pp，12~38。

因此，л 的有效区间为：

$$\frac{1-P}{P}\cdot\frac{1}{Y}>л>\frac{1}{1+X}$$

如果在二者之间没有空间，那就完全没有必要实施威胁。

$$\frac{1-P}{PY}\qquad\qquad\frac{1}{1+X}$$

或

$$\frac{P}{1-P}>\frac{X+1}{Y}$$

只有当 л小 于 1 时，才有必要实施威胁。

$$\frac{1-P}{PY}<1$$

或

$$\frac{P}{1-P}>\frac{1}{Y}$$

在这个案例中，犹豫不决的威胁优于坚定不移的威胁，而且这里根本没有必要做出一成不变的威胁承诺。问题的焦点在于威胁存在失败的风险，并且这一风险不受假设变量 л 的影响。这一假设确实有些特殊。如果我们假设 P 代表我们可能误解对手或夸大其选择右下格的几率，那么这一假设意味着人口收益呈现双峰分布。同时，这也意味着矩阵中的数字充分表明了一方的收益或另一个人的收益差距如此之大，以至于没有任何威胁能够消除这一差别。相反，如果我们假设右列与左列的结果比例表明人口收益显现振铃形分布，以及我们的对手是随机产生的，那么，我们的威胁成功几率将随着 л 的波动而波动。在众多偷盗犯中，被拘捕或被判刑的几率随着后者的几率而波动。我们上面分析的偷盗者可被分为两类：一类是为了偷盗钱财，一定会被矩阵中的收益震惊；另一类是为了一时好玩，他们则不会受到右下格的影响。此外，如果我们失败的几率反映了与对方沟通的失败，那么，我们完全可以做出这样的结论，即失败的几率不受某个成功威胁的影响。

如果认真分析上面的案例，你会发现我们将失败的几率与威胁联系起来，无异于直接降低威胁的效果。例如，假设右下格的 X 是对双方的惩罚，或者威胁如果被实施，那么，双方都将受到鞭笞或被关进牢

房。如果 X 代表行可以威胁使用的鞭笞和关押的最大期限，那么，π将代表行能够惩罚的最大有效数字。假设 π 为 0.5，则双方将均等分享应有的惩罚。如果我们以这种方式释义收益矩阵，并质问 π 的值多大时对行最有利，那么，我们就能以同样的分析方式得到同样的结论，即 π 的最小值为 1/(1+X)。因此，我们可以认为 π 表示威胁实施的几率，或一定能够实施的威胁程度。由于两个公式得到同样的结果，我们完全可以说，随机化在类似条件下的作用是在博弈选手之间分配一个规模过大、具有不可分割性的威胁，或分配一个规模更小的威胁。（需要指出的是，通过降低实施威胁的几率来削弱威胁，无疑相应减少了收益对双方的预期价值，而且直接减少的程度并非一定要与威胁的效果和价值成比例。）[1]

威胁意外实施的风险

还有一个“成本”因素可以降低威胁产生的几率。这个因素就是，威胁方可能无意识中实施了威胁，即使对方已经屈服于威胁（或者在威胁实施之前接受威胁）。同样，在盗窃犯或受害人屈服之前，威胁他们的手枪可能会走火。防范侵犯（他人土地）者的看门狗也能误伤无辜的人。

假设一个搭便车的旅行者用枪指着一个正在行进中的汽车司机，那么，司机可能以全速行驶、人为制造交通事故来威胁前者把枪丢到车窗外。在这种情况下，在搭便车的人充分理解司机的意图并接受其条件之前，完全可能提前发生交通事故。由此可见，意外履行威胁的危险也是威胁不可分割的一部分。然而，只有将威胁付诸实施才能使威胁具有说服力，因为除非司机全速行驶，否则搭车人是不会相信司机的威胁的。问题是，一旦司机将速度提起来后，在搭车人屈服和司机减速之间存在一个非常短的时间差，而这个时间差正是危险的高发期。当然，在这段时间内，高速度产生的危险或许能够被司机所控制。如果汽车速度一直

[1] 随机化或许是威胁计划不可分割的一部分，或者与决策过程有着千丝万缕的联系，无论威胁方愿意与否。因此，认为随机化是控制威胁程度的手段的观点仅适应于某些特殊场景。

保持在60公里以下，60公里则是交通事故的临界点，那么风险系数就不存在阶段性；在这种情况下，司机就不会轻易将速度提升到60公里。当然，如果搭车的人知道这一情况，那他也就不会相信司机的威胁。这样就存在“边际威胁”（fractional threat）的问题，即一方只是口头上进行威胁，却没有采取威胁的实际行为，从而不会造成人员的伤亡。但是为了达到威胁的效果，司机还得必须将这一切暂置脑后。

在这种情况下，我们假设——搭车案例基本上证实了这一点的正确性——意外实施的危险与威胁实施的几率л成正比（如果看门狗咬伤无辜行人的几率与咬伤侵犯者的几率成正比），那么我们将得到一个与前面公式略有不同的新公式。为了帮助大家理解，我们以同样的收益矩阵（这次我们忽视威胁失败的可能性）表示这一公式。假设л代表意外实施的可能性，其最小值保持不变。如果行希望进行威胁，那么他的预期结果必须大于0，公式表示如下：

$$(1-aл)-aлY>0$$

或

$$\frac{1}{a(1+Y)}>л>\frac{1}{1+X}$$

最理想的威胁必须稍微高于最低水平，最高水平л应该小于1。而且，这一威胁取决于X和Y以及参数a的值，或许л不存在有利的值。

随机性承诺

找到了边际威胁的规律后，现在我们探讨模棱两可的无条件承诺在某种情况下是否也是利大于弊。正如第3章和第5章所论述的，完全承诺——完全战略承诺等同于2人、2次选择博弈的第一选择。在这类博弈中，一方必须做出2次选择，完全承诺成为获得第一选择的手段。如果我们假设行拥有第二选择权，但是必须提前做出选择，选择i和ii几率各为50%，那么，我们就必须放宽这一释义。要做到这一点，一方必须保留进行第二次选择的权利，并利用这一权利提前做出选择。如果一方必须首先选择，那么他将失去随机承诺的机会。（选手设置的随机机制决定了随机承诺等同于“第一选择权”，但对方可能在其做出选择之

前尚未了解情况。）

如果我们对游戏规则进行适度修改，允许行在列做出选择之前做出无条件承诺，而且不能受到列的影响，那么，上面的收益矩阵（图 7–1）同样适用于这一新情景。如果行选择 ii，那么列很有希望选择 I。但是，这一选择对其没有任何好处，因左下格——行现在的承诺——没有任何回报。行的问题在于其不仅要以 ii 引诱列的选择 I，而且还要以 i 从 I 中获益。随机承诺也可能产生折中方案——承诺做出随机选择。如果在列做出选择后，行掷币选择（几率各为 50%）i 和 ii，那么，只要 X 大于 1，列就会选择 I。[1]那样的话，行将得到预期的 0.5。如果行将 π 的值设置成大于 1/(1+X)，那么只要列选择 I，行将得到最大预期收益。（如果列的收益是左下格而不是 0，即 0.5 或 –0.5，那么得到 π 最大值的公式也会不同。）如果行的左下格收益为 –1，那么他选择 ii 的几率不会大于 50%。假设这一收益是 –X 或更糟糕，就不可能出现 i 和 ii 的组合。如果 π 的值大到足以引诱列选择 I，那么，任何结果都是对行有利。

这是部分承诺的又一个规律。正如前面所讨论的，行更倾向于右上格的 I，从而将 π 的值减小到最小。在图 7–2 中，列期望行选择 i，即 π 的部分值。这样的话，行选择 ii 的承诺能够引诱列选择 II；行选择 i 的承诺将引诱列选择 I。无论行怎么承诺，列都不愿意选择 II。行无法以选择 i 来威胁利诱列选择 I，除非行允诺放弃选择 ii。在这些“完全战略”情景中，行都以收益 2 来结束游戏。如果行能够做出混合承诺（mixed commitment），其收益将会更好。他完全能够做到这一点，因为行和列同样倾向于选择 I，分歧之处在于双方对行的不同倾向。如果他选择 i 和 ii 的几率各为 50%，那么，列很有可能在第一列中的 2 和第二列中的 1.5 中选择前者。在这种情况下，行将得到 2.5，因为行倾向于 ii，虽然行在选择最佳行的同时希望列选择 I。因此，行希望的 π 值为：

[1] 即：只要列在右下格的结果小于左上格的结果或右上格的结果大于左上格的结果。详见前面有关结果比例的脚注。

$$4(1-\pi) > (1-\pi) + 2\pi$$

或 $$3/5 > \pi$$

这一混合承诺可以被称为边际威胁和部分允诺的集合体。实际上，行可以在列选择 II 时以 i 威胁列的同时，在列选择 I 时做出允诺。

如果行能够把 π 的值与列的选择相挂钩，行的收益将更好。假设行在列选择 II 时肯定以 i 威胁列，那么列将选择 I ，行选择 ii ，行将得到 0.75。但是，行做出威胁承诺的收益不能低于允诺的收益——如果行做出二者选择的几率相同——若 π 的有效值不能大于 0.6，行将得到预期收益为 2.6 （列得到的值为 1.6）。由于 π 的值不同，行做出允诺时，他得到的预期收益为 2.75（列的结果仅为 1.0）。

第8章

产生随机因素的威慑

战略威慑的一个显著特征是，实施惩罚性行为——如果威胁失败或必须实施时——对双方而言都是一个代价高昂、棘手的过程。虽然威慑注重的是事前预防，而非事后报复，但是为了让威胁更加可信有效，威胁方必须采取某些必要的威胁行为或为自己寻求实施威胁的充分理由和依据。人们对于美国在欧洲驻军，建立围堵前苏联人的铁丝网存在一种普遍的看法，认为这是美国为了让前苏联人相信一旦欧洲发生战争，美国不会坐视不管；不管前苏联人是否认为美国应该介入，美国绝不会违背这一承诺。

一般来说，如果对方不屈服的话，威胁方做出威胁承诺时必须表现出自己实施威胁的决心，并不是实施威胁的可能性。人们完全可以这么理解，如果威胁方表示自己可能实施威胁，那就意味着他可能不采取威胁行为。这也意味着威胁方给自己留下其他选择的余地——表明他没有实施威胁的决心。同样，威胁方表示可能实施威胁也容易使对手产生错觉；后者可能猜测对方是否会如实惩罚威胁方或自己，或选择放弃。而且，如果威胁方仅仅表示可能实施威胁——缺乏实施威胁的决心和意志——对手也会容易忽视威胁方的威胁，认为威胁方并没有实施威胁的意志。对手心里肯定会猜疑，一旦威胁方能够在实施威胁和退却之间做出选择，威胁方则会选择后者。（威胁方或许会自我安慰说他并没有撤

谎，因为他说的都是真实情况。）

尽管具有如此多的弱点和弊端，但是这类威胁行为还是存在可取之处。不过，只有当威慑实施过程的复杂程度胜过完全承诺的实施过程时，这类威胁行为才能取得实效；而且，威胁方也可能由于无意中实施了这类威胁行为，进而迫于情势不得不采取相应的行为，从而为此付出意想不到的代价。正因如此，这类威胁通常不为人们所认同和理解。

无论被威胁方是否拒绝屈服或威胁方是否采取威胁行为，类似威胁成立的前提条件是，最终的选择权并非掌握在威胁方手中。根据这一观点，真正的威胁不应该是“我可以实施威胁，也可以不实施，最终的选择权在我手中”，而应该是“我可能实施威胁，也可能不实施，但是最后的情况我也无法确定”。

是什么因素导致决策过程中出现不确定性呢？不管答案如何，有一点是肯定的：这个因素一定来自于威胁方的掌控范围之外。无论我们怎么称呼这个因素——机会、意外、第三方的影响、决策机制的缺陷或是一个我们一无所知的过程，威胁双方一定无法控制这一因素。一个典型示例就是意外战争。

意外战争威胁

很长时期以来，一直有人认为，很大程度上是偶然因素导致了一般战争的爆发，如偶发事件、错误的信号、机械失误、某个历史人物的恐慌心理、疯狂以及恶作剧，或者敌对双方对彼此信息的错误解读。一般来说，在某个特定的时期，当局势紧张或战略部队进入战争状态时，或者对对方的先发制人极度警惕时，敌对双方总希望将战争的可能性降到最低点。这时对于敌对双方而言，保持头脑冷静，避免冲动的决策、错误的判断和容易引起误解的行为就显得格外重要。无论是人为因素还是机械原因，意外战争似乎都根源于危机。

难道威慑威胁不属于这一机制吗？假设前苏联人注意到，每当他们采取进攻行为引起局势紧张时，美国就会迅速进入高度警戒状态，准备随时进行反击；假设前苏联人相信他们自己的立场——不断强调我们的报复性，那么，对他们来说，无形之中这些都增加了意外事件或错误信

息产生的几率。总之，不管是他们自身因素或是其他方面的因素，这些意外事件都可能导致战争的爆发。或许有人会问，难道他们没有认识到，全面战争能否爆发取决于他们自己的行为？如果他们采取进攻性或胁迫式行为，或者他们无法缓和对其他国家施加的压力，都有可能导致全面战争的爆发。

请注意，这里的问题不是美国决定发动全面战争，而是不管美国人愿意与否，战争一定会爆发。尽管前苏联人并不愿意对他们认为的不当行为进行报复，他们仍然会担心自己的行为可能导致一般战争的爆发，或激起以大规模战争或前苏联的大规模退却为结局的动态过程。前苏联人或许不会相信通过双方的共同努力，双方可以预测危机事态的后果，从而控制形势的发展。

这就是威胁——如果这一机制存在的话——我们可能采取大规模行动，请注意，不是我们一定这么做。这一威胁具有最高的可信度，其可信度来自于我们卷入针对前苏联侵略行为的大规模战争的可能性，而不仅仅局限于我们主动进攻前苏联这一事实。它扩展到了任何更加有效的威胁所能覆盖的范围。当前苏联对我们实施某种程度的侵略行径时，这就不取决于我们是否愿意或我们已经承诺进行大规模战争，而且最终的结果也具有很大的随机性。当然，如果前苏联人能够准确估计其行为的后果，那么我们或许可以避免卷入战争。

这样的威慑——假设我们将这一偶发行为机制称为威慑——向我们呈现了一些有趣的特征。不管我们是否意识到，威胁可能无时不在。甚至那些在过去数年里一直怀疑我们的大规模报复威胁能否威慑阻止小规模的侵犯行为，并为前苏联人很少做出错误的举动所迷惑的人也注意到，是前苏联人的行为导致我们采取类似的威慑行为，尽管我们并不愿这么做。而且，尽管我们不愿意发生意外战争并主动利用这一机制，但是我们采取的一些行为也可能导致这一威慑的产生。不管我们愿意与否，当我们（或前苏联人）针对危机事件采取相应行为时，双方都会遇到类似的威胁问题。认识了这一点，前苏联人也许会考虑风险因素。最后，尽管前苏联人没有发动战争就达到了自己的预期目标，但是这并不意味着威胁的消除。如果前苏联人制造危机后认为在未来的某个月里，

意外战争的几率会进一步降低，那么，他们将会得寸进尺，咄咄逼人。如果这样也不会发生大规模战争，他们仍然会坚信自己当初预期判断的正确性，并认为同样的举动也不会存在较大风险。前苏联人轮盘游戏中的任何幸存者都会认为，这样的行为没有丝毫危险。

有限战争：风险的根源

作为遏制侵略行为的有效方式，有限战争完全可以被理解为可能引发大规模战争的导火索。假设我们质问西欧人，当前苏联人对他们发动进攻时，他们该怎么做或如何抵抗前苏联人的进攻，那么答案通常涉及一系列决策：如果进攻规模有限，我们将决定打一场有限战争，而绝不会与他们拼得鱼死网破；如果我们能够以小规模战争抵抗前苏联人的进攻，那么前苏联人或许会放弃以前的想法，或者不会得寸进尺；如果这样，有限战争也许不会演变为一般意义上的战争，我们希望通过这种方式安慰前苏联人。即使不可能做出一系列显著可见的决策，至少在一方面非常显著：即人为的决策——选择是否主动出击，或者是否卷入战争，或者是否将战争升级，以及是否接受挑战。

对有限战争还存在另一种解释，即有限战争的爆发和升级肯定会增加全面战争爆发的可能性。有限战争可能导致两个风险：一方面，战争一方可能威胁直接将战争的成本——死伤人数、费用、领土损失、尊严的丧失等等——强加于对方；另一方面，战争一方可能威胁将有限战争发展为一般战争。[1]

这里，全面战争的爆发也是一种威胁。当然，这并不是说，一方采取某种行为将一定导致全面战争的爆发。在此我们必须再次强调，无论全面战争爆发与否，这都不是威胁方单方面所能控制的。正如谁也不能确定全面战争如何爆发——如同怎么产生误解、动机或过错，不管是什么因素导致两个大国间的有限战争具有巨大的风险，这种风险不是一方所能够消除的现实。因此，引发这一不可逆转过程的最终决策或关键行

[1] 格伦·H.辛德（Glenn H Snyder）在其文章中也强调了这一点。“拒绝与惩罚产生威慑”〔“Deterrence By Denial and Punishment”（Research Monograph No.1：Princeton University Center of International Studies，January 2，1959），pp.12，29〕。

为不应当被看做是双方的人为举动。“机会”或者是对有限战争本质和一般战争爆发的语境的不同判断，将决定一般战争是否爆发。

为什么一方愿意以有限战争而非全面战争来遏制对方的进攻呢？首先，以有限战争为威胁条件——根据分析得出的结论——是为了警告对方有限战争可能演变为一般战争，而非一定会出现这种情况。结果，这样的威胁产生的效果比大规模报复威胁的效果相对较弱，而且更适合于某种紧急事态。其次，当对方对我们的意图和承诺产生误解时，这样的威胁能够发挥缓冲紧张局势的作用：在能够避免爆发一般战争的范围内，双方都愿意卷入有限战争，进行双方可以承受的冒险举动。这样，我们不需要付出一般战争的价值，对方也可以通过撤军或停止进攻来降低风险。再次，如果对方缺乏理智或充满冲动，或者说我们误解了对方的意图和承诺，或者对方的挑衅行为程度超出了对方控制能力的范围，那么，我们在考虑是否实施威慑时一定要审慎，不能过于武断。如果我们认为以全面战争威胁阻止对方还为时不晚，那么，我们只能做出全面战争的威胁，否则就会降低威胁的可信度。假设我们以 1/20 的几率的全面战争威胁阻止对方继续前进，那么，如果对方继续一意孤行，我们则只能忍气吞声以 19/20 的几率妥协让步，从而避免一般战争的爆发。当然，我们降低自己风险的同时，也降低了对方的风险，尽管过多的安全措施也会削弱威胁的效果。但是，一旦我们完全误解了对方的意图，或错误判断了对方对其代理人、盟国或军事指挥人员的控制力，那么，更多适度风险将威胁到对方的全部控制力。

如果我们能以这种方式理解有限战争，那么，我们也可以对战争升级或战争升级的威胁进行合适的解释。根据我们的观点，我们不能根据短时间内的军事或政治优势来判断新式武器引入有限战争后产生的威慑效果，而只能以新式武器能否导致有限战争的升级作为判断的标准。正如在未来的三十年里，某个重要的因素可能导致有限战争升级为大规模战争一样，从常规武器到新式武器的发展将产生其他某个重要因素。

我们正是以这种方式解释“铁丝网”的存在。根据这一观点，我们向欧洲派驻军队毫无疑问意味着全面战争的可能性，如果一切正常的话；反之，则不能。我们现在拥有的是一系列的铁丝网，每一个都与随

机机制相联系；当对方从一个铁丝网向另一个铁丝网移动时，其意义愈加深刻。这里有必要强调，这一类比的重要意义在于铁丝网能否代表一般战争——至少一定程度上——是我们所无法掌控的，以及前苏联人是否知道这一点。

我们可以以同样的方式理解金门问题。有人认为，无论为之付出多大代价，是赢或是输，中国人和前苏联人只能被大规模战争所遏制，有限战争无法做到这一点。尽管他们相信我们已经使出浑身解数将战争控制在有限战争范围内，而且他们也为此付出了努力，但是他们仍然会简单地认为走向大规模战争的趋向并非双方所能控制或预见的。全面战争不爆发则已，一旦爆发，后果将不堪设想。

有限战争中的危险行为

如果有限战争的主要作用之一，是为了增加全面战争的风险，从而达到威慑对方并迫使对方以可承受的风险达到有限目标，那么，对有限战争行为的现有理解应该得到修正。或许有限战争的最高宗旨不是为了将战争维持在有限范围内，而是将全面战争爆发的风险控制在零点以上（*above zero*）的适度范围内。至少，这是一方“输掉”有限战争后能够选择的一种战略或者通过地区性的有限抵抗，甚至是两败俱伤的行为遏制对方的进攻态势。（同样，对方的进攻方式准备得越充分，以致地区性抵抗充满毁灭性的爆炸举动，那么，地区性抵抗就愈显无力。）

人为地增加全面战争的风险系数，是一种适合于有限战争语境的战术。当然，一方不可能仅仅通过口头宣传即可达到目标。一方不能简单地说：“昨天我们已经做好了全面战争准备工作的2%，今天达到了7%。你们最好小心点。”他应该采取行动——假设双方都愿意并希望将战争控制在有限范围内——让对方相信战争可能发展到双方无法控制的地步。

简而言之，这一观点认为有限战争的失控是一个逐步发展的过程。在战争的任何阶段，双方都能感觉到战争失控的程度。各种行为——创新、突破界限、不负责任的表现、挑衅行为、采取危险的战略举动、寻求稳定的联盟、欺诈和破坏行为、新式武器的应用、部队使命或冲突领

域的扩大——都会增强一方对形势失控的危机感。为了与对方分享同样的风险，一方也许会全力退出。一方最好通过不可更改的行为或承诺与对方分担风险，从而造成只有对方妥协才能实现和平的局面；否则，形势将发展为一场心理战。

报复与侵扰

有限战争并非人为的危险行为能够成为一种威胁的惟一场景。在大规模报复行为和有限战争之间还存在一个规模适中、逐渐发展的报复行为。一些有关有限报复战争的研究报告已经公开发表。[1]有人认为，如果前苏联军队进攻某个国家，应该有国家进攻前苏联的某个城市，并占领这个城市直到前苏联撤军。这一观点很大程度上虽受到了媒体的关注，但缺乏系统的研究分析。类似的观点还有主张进行小规模敌对行动——击沉战舰、封锁码头、制造沟通障碍等其他方式。

前苏联人的许多侵犯性和敌对行为既非足以引起有限战争，也非足以引起大规模的报复行为。这些行为大多涉及对对方的侵扰、恐吓、封锁中立国家或美国的盟友、以和平方式干扰我们的预警和雷达工作、将核武器视为心理战的一部分、怂恿对北约国家的各种敌对活动、支持各种叛乱活动，甚至使用非常规手段平息其卫星国的骚乱。我们采取行动应对这些行为并没有什么好处，更不用说我们卷入大规模报复行动，这更非什么明智之举。如果非要做出某种反应的话，可以考虑打一场一般战争，这既可以与对方共享风险，又可以避免巨大的战争成本。（如果这样也不行的话，至少应该将前苏联的不合理行为的目标和意义理解为妄图以一般战争、共担战争风险来恐吓对方。）

那么，我们该如何理解一个夸张的行为，如对敌对领土进行有限核报复？正如在有限战争中那样，对对方而言，成本包括两个方面：一方面是直接成本，如人员伤亡、破坏损失、尊严的丧失等等；另一方面是爆发全面战争的危险。没有人能够想像，如果一个国家在敌国投下一枚

[1] 最近，开普兰对此问题进行了深入探讨，参见“The Strategy of Limited Retaliation”(Policy Memorandum 19 of the Center of International Studies;Princeton，April 9，1959)。

原子弹会出现什么后果。如果这一行为被看做是一个孤立行为，意图有限，而不是故意进行的大规模袭击或对具有报复能力的敌国的突袭，那么，受害国如果大动干戈，不惜以全面战争进行报复，无疑是十分不明智的举动。当然，我们必须看到，即使被袭击国没有采取那样的行为，它也可能进行相应的、最终可能导致一场全面战争的报复行动。如果受害国以同样的方式进行反击，那么紧张局势或被缓和，或者继续升级。因此我们可以说，即使双方都竭力采取谨慎态度，但是双方对彼此意图的完全误解则可能产生一个最终引向全面战争的动态过程。

时机有时也可能对其十分不利。在此，我们探讨了什么行为才能导致或避免一般战争，敌对双方谁也无法控制最终的结果，全面战争的可能也只是一个主观判断的问题。我们提到这些行为的可能性并非意味着我们赞成这些行为，而只是向读者解释人们通常是如何理解这些行为的。对受害国的封锁禁运——威胁方也要为此担负相应的成本——亦有可能导致全面战争的爆发。

风险行为和“强制性”威胁

促使对方采取某种行为的威胁和阻止对方采取某种行为的威胁之间，存在显著的区别，即哪一方率先行动，哪一方的意志就首先受到考验。要想以威胁阻止对方的进攻，一方必须自断后路，背水一战。如果是想以威胁迫使对方撤军，那么一方必须表现出决死一战的气魄，不仅要自断后路、背水一战，还要顺借东风、火烧曹营。我可以将自己的车停在你的车道上，从而阻止你的车通过，尽管这一做法有些消极被动，主动权掌握在你的手中，撞不撞车由你决定。如果你发现我的车挡住了你的路，你可以以撞车威胁我离开，但是你同时也丧失了这一主动权。撞车的权利在你手中，然而我还是对你产生了威慑，继续阻止你通过。那样的话，情况似乎变得更加复杂。

迫使对方让步的威胁与阻止对方前进的威慑相比，前者更注重在对方采取某种行为后对其进行惩罚；而后者则更注重对方是否采取某种行为。之所以会出现这样的差异，主要原因是令某个行为具有可信度的惟一有效方式就是付出实施。尽管如此，在实施过程中，威胁方也可能受

到一定的损失，但是惟有威胁方采取了不可更改的威胁行为才能迫使对方屈服，从而减少双方遭受的更大损失，只有这样，威胁才能有效地发挥其应有作用。但是，如果双方都涉入其中的话，采取不可更改的破坏性行为或许并非最明智的选择。如果对方能够在短时间内做出妥协让步，从而将风险降到可以控制的范围内，威胁方实施风险适中的破坏性行为也许能够将威胁的风险降到威胁方可以承受的水平。威胁对方如果不屈服，那么，每天都将遭受大规模报复的可能性为 1%，在某种程度上等同于告诉对方（包括威胁方）每天将遭受总损失的 1%。（“某种程度”和“等同于”可以被灵活理解。）[1]

“摇船”，是一个好例子。如果我威胁你说：“快划，要不我就将船摇翻淹死我们。”你或许会说不相信。但是，如果我将船摇到快翻的程度，你就会相信了。即使我不能将船摇翻淹死我们，我只需要将船摇到快翻的程度，同样可以发挥有效作用。然而，要使威胁看起来更可信，我必须将船置于危险境地。换句话说，我威胁将船摇翻有点危言耸听。

为了实现这个目标，我应该有一个黑夹子，里面藏着一个能够挑起一场全面战争的转轮和一个仪器。然后，我就把夹子放下，告诉前苏联人我已经将这个夹子启动，并在未来的某天发动一场全面战争（提前设定参数，并告诉前苏联人）。我告诉他们——我能够向对方证明——这个夹子会一直运转到他们答应我的条件为止，我是没有办法让它停止的。请注意，我并没有强调自己能够决定全面战争的爆发或人为地发动，如果这个夹子触及某个参数的时候。当某个正确（错误）的组合某天来临的时候，我任由这个夹子自动将双方一起吞噬。[2]

尽管对方已经做出了妥协，但是在他平心静气做出正确选择之前，只要这个能够引发战争的夹子继续存在，只要我没有确定夹子在某天一

[1] 如果一方不能以坚强的意志进行冒险行为，那么，任何冒险行为都不会对对方产生威慑效果：后者依然相信能够以坚决的行为迫使前者退让。假使后者只是一时的全力以赴，那么，前者仍必须赢得“心理战”的胜利，因为这一对称情景至少取代了对对方有利的非对称情景。在后一种情景中，如果双方都不采取某种行动，对方将自动获胜。

[2] 这一策略风险越小，机制的自动化程度越强。而机制的自动化程度越强，敌人在心理战中希望得到我方意图的动机就越少，风险持续的时间就越长。

定会爆炸，那么我就仍然还掌握主动权。在普通的威慑中——只要对方按照我们的要求行动，什么事都不会发生——如果以过高的条件提出威胁承诺，那么代价将过于高昂，而非出于自卫。在目前情况下——正如我们所预期的，威胁随着时间的流逝而逐渐增加——一个条件过高的威胁将无法实现预期目标。在这种情况下，一个可能性很小的威胁不再仅仅是可能性很大的威胁的潜在替代品，而成为后者的最佳和必要选择。例如，一个拥有中等规模核力量的欧洲国家让前苏联人滚出匈牙利，否则其将对后者进行大规模进攻。毫无疑问，前苏联人对这个威胁将置之不理，因为前者无法做出实际行动让前苏联人相信它具有这种能力；相反，假设这个欧洲国家威胁向前苏联一天发射一颗携带核弹头的导弹，或者发射某天在前苏联境内自动爆炸的装置，如果前苏联人没有将其击落的话。前苏联人或许会说，他们不相信对方会这么做。但是这个国家却做到了，前苏联人对此发出强烈抗议并进行威胁。一天过去了，这个国家又继续向前苏联发射了一颗同样的导弹，或许通过了前苏联的防守，或者几个弹头穿过，或者一个也没有。假设一些导弹躲开了前苏联人的拦截，在前苏联的一些城市中，或者人口密集区，或者荒芜地带爆炸，那么这个国家还会继续这样做下去。

这个国家究竟什么时候才肯罢休？这个国家想要达到的一个重要目标——除了给前苏联造成损失，打击前苏联的气焰外——是为前苏联和自己（也涉及全世界）创造一个可能在不久的将来引发全面战争的危险，尽管双方都不希望出现这样的战争。这个国家可以理直气壮地对前苏联人说："如果你不从匈牙利撤军，我们就打一场全面战争吧！"前苏联人什么时候会全面妥协呢？前苏联人越早撤出，全面战争的危险消亡得越早。为了同样的目标，这个国家并没有这么说："赶快从匈牙利撤出，否则我就发动一场战争。"这个决定由不得他们决定，也不取决于他们是否表现出做出最终选择的决心。前苏联人猜测对方可能会竭尽全力阻止全面战争的爆发，但是同时他们也必须承认，只要这些导弹不断地飞来飞去，时不时地爆炸，无论他们如何相应地反应，双方谁也不知道如何阻止全面战争的爆发。

这个案例意在通过喻例说明，产生全面战争危险的行动作为形势发

展的一部分很难被战争双方所觉察。以最近发生的事情为例，假设在地面部队无法接触柏林时，我们派一个飞机纵队；或者假设一旦柏林的交通压力令人无法忍受，必须有派往地面部队建立交通走廊；或者假设无论有意无意，我们采取的行动导致了东德的起义，那么我们如何分析这些事件对前苏联人的压力？我认为答案很大程度上是，他们可能发动一场双方都不希望看到的战争。尽管双方都不愿意发生战争，但是二者对此还是无能为力。一个直接行动的方法——甚至小规模行动也可能达到部分目标——就是人为创造与前苏联人共担的风险，迫使他们在屈服、撤退和回击中做出选择。

当然，这并非对类似行动的惟一解释。假设我们打一场小规模战争，我们也有可能取得军事上的胜利。如果前苏联人想要将战争升级，则必须采取破釜沉舟的战略，因为只有这样才能避免我们的强烈反应。在这种情况下，有限战争的初期阶段将存在一个阻止战争升级的“威慑”威胁。尽管如此，能够解释为什么小规模战争威胁也能发挥作用的一个重要依据是，小规模战争威胁为敌对双方提供小规模战争场景的同时，增加了大规模战争爆发的可能性。这一可能性小则足以令前苏联人相信西欧国家会发动这样的战争，而大则足以令双方两败俱伤，假设战争爆发的话。[1]

假设我们没有取胜的绝对把握，那么，有限战争威慑能否发挥有效作用对我们来说没有任何意义。根据这一解释，一场有限的地区战争可能并非仅仅是地区性的军事行动；同时，这一解释暗含了可能对前苏联本土进行的报复行动——不是有限的报复行动，而是大规模报复战争。

边缘政策

本书的观点涉及到“边缘政策”的定义和“战争边缘”的概念。按照本书的观点，这里的边缘并非如悬崖绝壁那样，一个人可以牢稳地站在上面四处张望，冷静地决定是否跳下去。这里的边缘带有一个曲滑的

[1] 在作者看来，美国 1958 年向黎巴嫩派军既是冒险行为，也是一次成功的行动。成功取决于风险——共产主义国家的反应决定了风险的大小。

斜面，一个人站在上面，一不小心就有掉下去的危险；距离边缘越近，就越危险，掉进悬崖的可能性就越大。但是，无论是站在悬崖上的人还是旁观者都不知道危险究竟有多大，也不知道愈靠近边缘时，危险大到什么程度。实施边缘政策并非意味着，一方为了威胁围攻自己的对手而故意不断接近边缘，从而当自己想跳下去时，别人拦也拦不住。边缘政策意味着一方不断接近边缘，从而可能与敌人同归于尽，即使自己想自救也为时已晚。[1]

由此可见，边缘政策是人为制造的战争危险，而且一旦出现就必然失控。边缘政策是一方人为造成形势失控的策略，因为只有形势失控才能迫使对方屈服妥协。这意味着通过故意向对方暴露共担风险，达到侵扰和威慑对方的目的；或者警告对方，如果其做出相反行为侵扰我们，我们就可能拉他一起跌进危险的深渊，不管我们愿意不愿意，从而阻止对方继续前进。

有人认为，我们应该始终让敌人猜测自己的意图，特别是我们是否会做出反应。现在看来，这一观点需要重新诠释。也有人认为，我们没有必要以坚决的报复或抵抗来威胁敌人，仅仅以可能性的反击即可达到同样的目标。但是，如果将可能性的反击用于前苏联人身上，那么，这一观点的说服力就显得多少有些苍白。前苏联人或许认为，我们受到进攻后不会进行必要的反击，特别是当他们发动中等规模的侵略战争时。假设我们不愿意进行必要的反击，前苏联人就更加认为我们一定不会回击。如果我们为自己留有余地的话，那么，我们似乎已经验证了对方对我们的预期判断。因此，如果我们担心明确的威慑承诺可能会失败，从而达不到最初的目标，结果做出一个似是而非的决策，给自己留下妥协的后路，那么，这无异于告诉敌人：当危险来临时，我们则可能违背自己做出的威慑承诺。

但是，假设我们能够使前苏联人明白，我们已经做好了充分准备，尽管我们有可能回避，那么，情形将大不一样。有人认为，我们不会对中立国的侵犯行为进行报复（取决于对方的进攻时机）；或者我们不会

[1] 儿童对这一点理解得更为透彻。

让前苏联人为自己拿主意，也不会让对方知道未来的结局。这种说法在前苏联人看来，完全是自欺欺人。派往部队进驻中立国或到中立国的附近地区，或做出一个我们也无法确定自己是否会在侵略出现时回避战争的承诺，也许会引起对方的猜疑。

总而言之，尽管主观上不愿意让敌人猜测我们的动机意图，但看起来我们似乎还是在让敌人不断地猜测。如果结果一定程度上取决于意外因素或者是一个我们显然无法控制的过程，那么，我们确实在为对手制造威胁。

不完全决策过程

强调威胁是一方“可能”进行报复或发动战争——决策过程失控——意味着政府的最重大决策往往是一个不能完全预测或人为能够“控制”的过程。同时，这一观点也暗示由于决策过程的不完全性，即危机发生时双方很难准确无误地预期判断对方的决策过程，从而产生误解并导致决策失误。一个国家也可能因此意外地卷入到一场战争中。当然，决策过程也可能存在其他随机因素，如信息错误、沟通不畅、误解、权力滥用、恐慌，以及人为失误或机械故障。

这一观点并没有反映人们通常对决策过程的理解。第一阶段，决策应该建立在非完全依据或模棱两可的预警信息基础之上。同时，原则上否定错误警报可能会导致不可更改的行为也是毫无道理的。（而且，人们不应该错误地认为，可能存在某个特定水平；在这个水平之下，特定的危险不会发生，除非更严重的危险出现。）

第二阶段，战争之所以能够发生是因为双方都进入决策的死胡同，无法进行协调，谁也不愿妥协让步，特别是当退让触及核心军事问题时。我们必须承认，国家政府之间也存在对彼此意图的错误判断。

但是在第三阶段，甚至是一个组织井然、负责、决策相对冷静的政府也可能存在不完全决策过程，特别是当危机发生时。尽管造成这一现象的原因众多，但是其中之一则是决策过程往往涉及完全中央独裁式的决策群体。他们在价值观、对对方意图以及军力的判断上都存在不同。在紧急状态下，决策往往取决于现有决策者、已经完成的某个研究课题

以及面对史无前例的挑战的某个领导者和智囊人物的主张和权威。当然，一部分决策也可能来自于总统或首相或内阁授权的代表机构和全权代表等；而且，如果前者的授权经过国会或议会的批准，那么，决策过程中涉及的人将会更多。在某种决策过程中，甚至存在相互冲突的地方，如宪法的规定与紧急决策之间的矛盾。因为在紧急状态中，决策人员需要就事论事做出临时的决策，而宪法规定已经存在，无法临时修改宪法来满足暂时的需要，从而存在违宪的可能。当出现这种情况时，解决的方式往往是默许违宪行为的存在，而非大张旗鼓地予以公开。此外，保密的需要也限制了紧急事态的准备工作。

正因如此，对于政府而言，不存在一个固定、普遍的决策模式以应对所有的紧急事态——即使是可能预见的意外事件。在未来的危机事件中，如何进行统筹决策，需要考虑什么利益因素以及集体决策机制如何运行，都无法提前简单地做出周密安排。

如果我们提前认识到处于战争边缘的普通人的智力极限和政府决策人员的情商极限，那么，我们将更容易理解为什么当出现这一情况时，国家似乎已经成功地超脱其身；虽然国家已经冒着巨大风险，在允许的范围内尽了最大努力，但是仍然存在失败的可能性。

我们不能期望政府唤起公众注意其在这方面的弊端，以及向敌人表露自己的不完全决策是战略的必要组成部分。从具有重要意义的公共关系角度讲，一方也不能将自己的判断失误和错误警告或者无法从危险形势中脱身透露给对方。一般情况下，我们也可以理解身处有限战争的政府为什么总是对外声称，自己采取军事行动并非出于对全面战争的考虑。问题在于这些行为都在暗中操作。

尽管在大多数情况下，我们都无意识地运用这一威慑策略，甚至有时是人为地运用，但是理解产生随机因素的威慑的基本含义无疑具有重要意义。在第一阶段，对方可能利用这一威慑策略对付我们；而在第二阶段，如果我们无法意识到全面战争因素的存在，我们也可能错误判断自己运用的策略。即使我们没有给予这些因素应有的重视，它们也是我们对对方施加影响的重要组成部分。假设——举例说明——这是欧洲国家拥有的有限战争力量的重要组成部分，那么，我

们对这一作用的分析可能会引起人们的误解，如果我们没有认识到这一点的话。通常认为“铁丝网”要么发挥作用要么无效，或者前苏联人希望“铁丝网”发挥作用或无效的观点，都错误地将愈加复杂的可能性简单极端化。

第四部分

突袭：一种互不信任的分析

第 9 章　双方对突袭的担心
第 10 章　突袭与裁军

第9章

双方对突袭的担心

假设某天晚上，我听到楼下好像有什么动静，就拿起枪下楼看看发生了什么事。结果，我和一个正在作案的小偷正面相遇，而且对方手里也有一支枪。现在我们面临一个双方都不愿意看到的场面。尽管对方希望安静地走开，这也是我所希望看到的，但是我们之间还是存在一个危险：对方可能认为我会开枪而首先开枪。更严重的情况是对方可能认为我认为对方可能开枪，或者对方可能认为我认为对方认为我想开枪。当双方都希望在自卫中免遭枪击时，“自卫”将变得愈加模糊不定。

这是一个有关突袭的问题。如果突然袭击能够带来某种优势，那么，人们往往选择先发制人，采取突袭行动。由于担心我们的突然袭击，对方可能率先发动袭击；相反，我们也可能出于同样的担心，而率先发动突袭。但是，如果突袭成功带来的收益远远小于战争成本，那么双方就不存在发动突然袭击的“根本”动机。尽管如此，似乎还是存在某种特殊诱惑促使双方先发制人，对对方进行突袭——尽管突袭的动机如此之小——这种诱惑可能演化为一个复杂的混合过程，即包括双方的预期互动和一个特别的循环逻辑引起的袭击动机“他认为我们认为他认为我们认为……他认为我们认为他会进攻，因此他认为我们会进攻，因此，他会进攻；因此，我们必须进攻”。

有趣的是，尽管这一问题通常出现在我们平时所指的冲突场景中，

如前苏联人和我们之间的冲突或小偷和我之间的冲突，但是，这一问题更像两个或多个搭档之间缺乏信任的问题。如果每个人都想从共同财产中攫取自己不应该得到的部分，那么，每个人都会怀疑其他人也在谋划同样的事情。如果大家都知道别人也存有疑心，那么，他就会怀疑自己是对方的怀疑对象。我们现在设计了一个能够代表突袭问题的收益矩阵。如果焦点是暴民中的某几个人，那么，剩下的暴民可能将这几个人驱逐出去以防止他们率先揭发，而身处危险的人则可能出于自卫率先揭发。因此“预防性自卫”（preclusive self-defense）的博弈结构与“伙伴互信”（partnership confidence）的博弈结构相同。

有人认为，敌对双方间的不信任可能增加了突袭产生的几率——并可能产生多重效果——这也正是我在本章所要探讨的问题。具体说，我希望分析对这一可能性的理性思考或理性战略选择能否导致这一现象出现，如果双方选手充分认识了二者之间的困境的话。上述观点——或许我对其产生误解——可能是客观存在的现象或动机行为。人们可能模糊地认识到这一形势的敏感性，并可能人为地将其激化。但是我想得到的结论是，这一“混合预期”（compound expectations）现象能否被看做是一个理性的决策过程。我们能否建构一个明确的困境模型？在这个模型中，两名理性选手都将是各自预期判断的受害者。[1]

概率的无限排序

我们首先以下列步骤设计这个问题：假设一方选手控制一组概率，并可能存在无限排序。排序的第一项是预期目标 P1，这也是对方“希望”袭击的目标——即使他自己不担心受到袭击，对方也要对其发动袭击；第二项为预期目标 P2，这是对方认为我所“希望”袭击的目标；即使我不担心自己会受到袭击，也要对其发动袭击。第三个预期目标是 P3，这是对方认为我认为他“希望”袭击的目标；第四个预期目标是 P4，这是他认为我认为他认为我“希望”袭击的目标。第五个、第六个

[1] 博弈理论家将这一问题看做是零和博弈的对立面非零和博弈，即“二人博弈”。在这里，非零和博弈关注的是能否开枪，而非何时开枪。

等等依次排序。每个排序中都包含“他认为”和“我认为”两个量。结果，我们得到的对方可能发动袭击的总概率为：

$$1-(1-P1)(1-P2)(1-P3)$$

这个公式的缺陷在于没有产生级数（series）。每一个目标都是特定的假设，并反映了有关特殊场景中具体信息结构的数据。但是，我们无法仅仅以几个级数数据为基础推断此后数据的无限性，或者不管能推理到什么程度，以几个数据数理决定整个排序。排序的项数最多表示一个选手拥有充分时间进行预期判断，或者需要记住的最根本因素，因为他必须独立建造排序的每一项。我们也能够建构一些博弈游戏，并使其信息结构具有产生排序的功能——例如，转轮的一系列旋转能够决定另一方选手是否知道我的“真正”价值观；或者我是否知道他是否知道，或者他是否知道我已经知道他已经知道等——但是这些游戏也可能仅仅是特例，因而无法说明我们希望得到有关普遍现象的分析结论。我们需要的是问题的明确表达，能够允许我们随意使用变量，这些变量可能代表排序的首项或“目标”（objective）项。在某种情况下，这些变量可以自动赋予“他认为我认为”这一无限循环可能产生的每一项的值。我们必须建构一个博弈场景，在该场景中双方的预期判断为对方服务。

具有“明确答案”的非合作博弈

作为第一步，我们给每个选手赋予一个初始变量，表示他可能发动袭击，即使是在不应该的情况下。双方完全知道这些变量的值。我这里所指的“不应该的情况”存在于下面对双方行为的假设中。

行为假设的第一部分是，如果双方能够意识到互不进攻政策是对双方最有利的结果，那么，他们就会承认这一“解决方案”并主动放弃进攻。例如，假设图 9-1 表示了这个收益矩阵，选手双方都相信对方会做出对双方最有利的战略选择。对于选手理智的这一要求似乎并不高。[1]

[1] 按照卢斯和雷夫的观点，如果非合作博弈具有“严格意义上的解”，那么，这个“解”能够具有普遍性，见《博弈论与决策》，第 107 页。实际上，这一条件在这里占有很大优势，因为选手双方都倾向于这个解，不选择其他所有备选项，而不仅仅是平衡点。

（我认为，这是一个值得商榷的问题，主要原因是双方同时做出不进攻努力的可能性远远小于单方面突袭。正因如此，双方选手很难相信对方不会突袭。一方可能出于自卫或者担心对方的突袭而先发制人，从而破坏了这一原则。这一可能性存在于行为假设的第二部分。）

	Ⅰ	Ⅱ
ⅰ	0, 0	0.5, −0.5
ⅱ	−0.5, 0.5	1, 1

图 9-1

在假设的第二部分中，选手 R 可能袭击目标 Pr，选手 C 可能袭击目标 Pc。当二者能够（应该）选择非进攻战略时，他们可能进攻这些目标，即他们的决策违背了我们的第一个假设。这正验证了选手可能在他“不应该”的时候发动进攻。正如这一变量代表的：选手出现非理性的可能性；或收益矩阵自身出现缺陷的可能性；或选手选择单方面突袭的可能性；或选手出现决策失误，意外调兵遣将发起进攻。在我们的决策模型中，选手的变量参数都来自于外界：外界因素产生的数据，而不是两名选手自身创造的。

假设选手双方对 Pr 和 Pc 两个变量十分清楚，在他们之间不存在什么秘密或猜测，那么，这一假设似乎有助于我们解决问题，但是事实并非如此。这两个外界因素导致的袭击行为，表明了双方在现实中可能进攻的目标。问题在于经验，即存在不确定因素的情况下，双方的预期互动是否会产生新的袭击动机。因此，我们至少应该在游戏中设置预期判断和猜测可能需要的变量参数。将随机导入因素控制在最低水平内的惟一方式，是促使双方选手意识到这两个变量的存在；否则，我们必须把他们对两个变量的猜测表述清楚，如一方将猜测对方对两个变量的猜测，或一方猜测对方如何猜测自己对两个变量的猜测等。我们必须再次得出特殊情况的无限排序，即概率分布的格局这一难题。划清界线或划分双方对突袭担心的惟一方式是，使基本的不确定因素成为双方记录的

问题。我们希望得到的结论是，不确定因素的“客观”源是否是双方主观恐惧的“上层建筑”（superstructure）。

现在，我们面临的情况看起来似乎产生了我们所说的自卫情景。第一个选手必须考虑对方袭击的可能性，以及对方是否也像自己一样担心袭击。甚至是完全没有进攻意图的一方，也必须考虑到对方可能非理性地或出于过分担心遭受袭击而先发制人，首先发动进攻，看起来我们似乎得到了一个动机综合体。

但是，实际情况并非如此。我们并没有得到期望的任何常规乘数效应。双方发动袭击的可能性并没有增加袭击产生的几率，除非他们坚持一定发动袭击。尽管博弈源于双方对彼此发动袭击的无限可能性，但是博弈的结果是，双方对袭击的担心并没有增加袭击出现的可能性，出现同时进攻或者彻底消除突袭的可能性，即双方同时决策或者没有任何可能性。

在解决问题的过程中，我们借用了两个变量表示非理性袭击出现的概率，用于检验上面的矩阵结果。矩阵中的左上格数字仍然不变，右下格的收益被重新验算，作为 4 个格的平均值。结果，如果双方都选择不突袭，那么，最后的可能性结果是（1–Pc）（1–Pr），即没有袭击出现；如果结果是 Pr（1–Pc），那么，R 可能发动进攻，而 C 不会；如果结果是 Pc（1–Pr），那么，C 可能发动进攻，而 R 不会；如果结果为 PrPc，那么双方都有可能发动进攻。同样，左下格的结果是左列结果的平均值，因为如果 C 选择进攻，他一定会发动进攻；然而，如果 R 选择不进攻，那么当结果是 Pr 时，他可能选择进攻；但是当结果是（1–Pr）时，他也可能不进攻。那么，既然双方的非理性袭击系数均为 0.2，我们上面的矩形将得到一个新矩阵，如图 9–2 所示。[1] 如果 C 的非理性袭击系数为 0.8，R 的袭击系数为 0.2，那么我们将得到矩阵图 9–3。如果双方的非理性袭击系数为 0.8，那么我们将得到图 9–4。

[1] 实际上，我们认为选手是在“完全战略”和含有自变量的“混合战略”中做出选择——根据博弈的说法。（当然，他们可以进一步将完全战略和混合战略结合在一起，但是这里他们没有理由这么做。）

	I	II
i	0 0	−0.4 0.4
ii	0.4 −0.4	 0.64

图 9–2

	I	II
i	0 0	−0.1 0.1
ii	0.1 −0.1	0.46 0.14

图 9–3

	I	II
i	0 0	−0.1 0.1
ii	0.1 −0.1	0.04 0.04

图 9–4

在第一个矩阵图中，双方出现非理性袭击的可能性——即双方的袭击系数均为 0.2——是无伤大雅的，至少就战略选择而言如此。它们产生了一个新的收益矩阵，在该矩阵右下格存在一个“明确的解决方案”。鉴于无法回避这两种情况，所以，双方在游戏中的收益将得到相应的减少。但是，两个因素的存在也阻止了双方采取进攻行为。深入考虑这两个因素的选手都意识到，通过前面假设情形中选择的那个战略，在二者之间仍然存在和平的可能性。

最后一个新矩阵中，每一方的进攻系数均为 0.8。尽管双方行为存在对称性，但是这并不妨碍双方间存在不稳定因素。每一方都有可能发起对对方的进攻，而不是主动避免战争。双方都知道对方也会这么做。

如同博弈论中“囚徒困境”现象一样，这是一个普遍存在的现象。如果强制性协议在机制上具有可行性，以及游戏允许双方利用最后的机会达成一个类似的强制性协议，[1]那么，这个强制性协议将是双方避免战争的惟一有效方式。（尽管如此，双方在游戏中的收益都将减少 0.04）。

在第二个新矩阵中，尽管双方之间不存在对称性与否的问题，但是双方间还是存在不稳定因素。C 的非理性可能导致 R 出于自卫而先发制人，首先进攻。C 也可能出于相同的想法而采取同样的行动。[2]

[1] 博弈论中的“囚徒困境”主要涉及影响选手双方战略选择的结果格局——在缺乏强制协约的情况下。这个术语来源于两名被分别审讯的囚徒问题。这两个囚徒要么取得一致口径，争取得到较轻的惩罚；要么一方率先招供，争取宽大处理被释放，而另一方则被从严处理。详情请参阅卢斯和雷夫案例， pp.94ff。

[2] 如果我们假设 Pr 的值为 0.2，Pc 的值为 0.6，那么，我们得到一个略微不同，但十分有趣的场景。修改后的矩阵（仅供 R参考）如下所示：

0	0.2
–0.4	0.12

R 依然掌握进攻“战略的主动权”，即不管 C采取何种行为，主动进攻为 R 最有利。但是与图 9–1 场景不同的是，这里如果双方都不选择进攻战略，那么 R 将处于不利境地。C知道 R 的战略主动权可能导致双方得到 0。代表 C 的“非理性”的 Pc 使 R 出于自卫不得不选择进攻，而且其中—— 自卫动机“不纯” ——也会促使 R 发动突袭，从而赢得主动权，避免坐以待毙。假设 R在无法对 C 发动突袭的情况下强行进行突袭，那么，他在原有矩阵右上格的结果为 0，而不是 0.5，表示 R 的收益矩阵修改后如下所示：

0	0
–0.4	0.8

尽管右列的两个结果对 R都不利，但是相对而言，上面的结果略优于下面的结果。如果情况如此，那么 R将取消进攻 C 的动机，C 也知道这一情况，则双方的结果将是和平共处。不仅非理性选手的无法进攻对选手双方有利，而且“受害者”的无法“自卫”也有利于双方。这一情景能够以变量表示为下面的公式：

$$1-h<Pc<1/(1+h)$$

这一结论可以被广泛地应用到其他场景。假设 h 表示赢得战争得到的回报，且 h 大于 1。假设情况如此，而且进攻总是致胜战略，那么，当对方无法进攻时，双方的首要战略选择均为“主动进攻”。当双方为了得到更大的收益而采取逃避现实的策略时，双方的收益均为 0。现在，我们假设 Q表示突袭并取胜的可能性，那么，单方面突袭所得到的预期收益只能为 Qh。

由于 Pr 和 Pc 的取值超过一定限度，将导致双方之间存在不稳定因素，并引发战争，因此二者的取值极限必须为：

$$Pc<1-hr$$

$$Pr<1-hc$$ [1]

（假设 h 表示选手一方单方面获得的收益； -h 表示遭受袭击而没有采取反击的一方选手的损失；当双方同时发起进攻时，二者的收益均为 0；当双方能够避免战争时，二者的收效均为 1。）

图 9-5 表明了选手双方在游戏中的收益分配和每个战略选择具有的价值，即 P 在 0~1.0 之间移动。假设 Pr 的取值为 0.2，那么游戏形势将对 C 不利（与图 9-1 相比）。在这种情况下，C 和 R 的收效如图所示。当 Pc 的取值为 0.5 时，双方的关系将十分紧张，二者收益均为零。

当 Pc 和 Pr 的取值不等时，我们完全可以忽略取值较小的变量。但这并不表明，游戏完全不符合“混合概率”（compounded probability）这一宗旨。如果二者都低于关键点，那么，它们的取值将没有任何意义；反之，如果二者之一的取值高于关键点，另一个取值是 0 或 1.0 也同样没有意义。当超出某个极限时，它们仍然有效，意味着双方可能引发一场战争。因为，二者导致选手双方的战略选择从和平战略转向了进

如果 Qh 小于 1，那么，我们得到的矩阵是双方的最佳选择——和平共处。如果存在“非理性”进攻的可能性，那么当 Pc 小于（$1-Q_c h$）以及 Pr 小于（$1-Q_r h$）时，博弈场景最稳定。假设 Pc 和 Qc 满足其中第一个条件，则满足第二条件对 R 和 C 都有利。如果 Pr 不在考虑范围之内，那么，R 则希望 Qr（代表 R 突袭 C 的能力）不小于（1-Pr）/h。只有这样，R 和 C 的收益才能大于 0 如果 R 能够为对方升级预警系统提供资助或者以显著方式削弱自己的进攻能力，从而使 Qr 的值控制在范围之内，R 也能够达到上面的目的。这一原则同样适用于彼此不信任的二人选手——双方将各自的枷锁扣在对方的头上。如果一方无法提供枷锁，那么，对方将主动提供自己的枷锁。只有这样，双方才能实现双赢。

[1] 下面是一个表示非对称性博弈场景的更普遍公式。公式中的 R_{11}、R_{12}、R_{21} 和 R_{22} 分别代表 R 在行 1 列 1 和行 1 列 2 中的结果，依此类推。公式的分子表示在错误的时机发动进攻所付出的代价；而分母则表示错失进攻良机所付出的代价。二者的标准相同，或许表现为 P 和（1-P）为肯定概率而非二者偏离概率，而且符合“理性”行为标准。

$$\frac{Pc}{1-Pc}<\frac{R_{22}-R_{12}}{R_{11}-R_{21}}$$

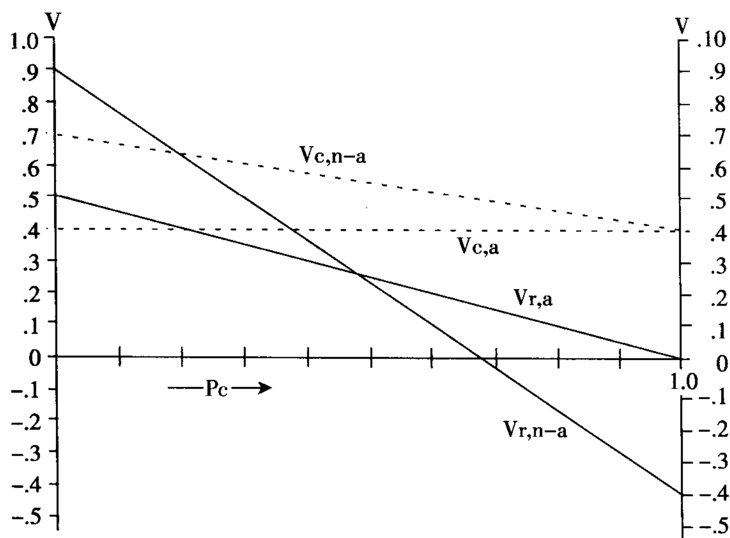


图 9-5

P_c 的取值决定了 R 和 C 在游戏中的收益； P_r=0.2。Vr，n-a [=0.9-1.3Pc] ：双方维持和平时，R 在游戏中的收益；Vr，a [=0.5-0.5Pc]：C 选择不进攻，R 进攻所得的收益；Vc，n-a [=0.7-0.3Pc] ：双方维持和平时， C 在游戏中的收益； Vc，a [=0.4]：R 选择不进攻时，C 进攻所得的收益。

攻战略。导致战争爆发的因素则可能来自于外界，也可能是选手的主观因素。

如果我们假设选手轮番选择，这并不影响我们得到同样的结论。假设 R 拥有选择进攻和不进攻的自由，而 C 则必须坐以待毙，那么，只有当 R 有机会做出选择并付诸实施，却没有发动进攻时， C 才能采取进攻行为。现在，我们进一步假设在 R 做出选择之前，C 拥有首先选择权。在 C 选择之后，R 做出选择，最后是 C 做出选择。然后，我们再给 R 一个优先选择权。当然， R 也可以首先选择，然后 C 做出选择，接着 R 做出选择，最后是 C 做出选择。（依此类推，直到有一方采取进攻行为为止。）

那么博弈结果如何呢？如果是在图 9-1 中，在做出最后一次选择时，C 将不会选择进攻；但是，如果 C 在此选择了进攻，那么他可能得

到变量 Pc。在做出最后一次选择时，R知道C会选择进攻，所以，事先针对 Pc 制定了相对对策。在做出上一次选择时，C 知道 R 将做出什么选择，所以，针对 Pr 做出了自己的对策。以此之前，R 预测到了 C 在未来的两个选择中可能选择进攻，所以针对 $1-(1-Pc)^{2}$ 事先做好了准备。依此类推，如果双方都有无穷次选择的机会，分别有 Pc 或 Pr 个概率对对方进行非理性攻击，结果将取决于 $Pc=1-(1-Pc)^{n}$ 和 $Pr=1-(1-Pr)^{n-1}$ 是否满足前面推理的条件。如果情况果真如此，选手双方都知道对方下一步不会选择进攻，那么，自己也不会主动进攻。但是如果 P 的取值大于某一方可以承受的程度，那么，他可能倾向于发起进攻，对方也会知道这一点。结果，无论谁首先做出选择，都会首先选择立即进攻。

换句话说，虽然我们将双方的预期判断进行了综合组合，但是我们取得了最大最小效果。而且，我们在组合过程中也没有要求双方选手必须将各自的非理性变量结合起来，因为两个或至少一个变量的取值和游戏的时间已经足以满足第一个选手发动进攻，或者双方保持和平共处。如果我们假设所有非理性变量不受轮次的影响，而且以 $1-(1-P)^{1/n}$ 表示每轮的变量，那么，这些组合变量将为 Pr 和 Pc，游戏的结果则不受轮次的影响。如果我们将这个游戏看做是“他认为我认为”情景，每一轮选择都代表每一轮的怀疑过程，那么，我们将得到一个新模型。在这个模型中，双方依旧持续不断地相互怀疑，或者为各自选择进攻提供“客观”条件，或者双方都放弃进攻。

重温旧题

现在，我们发现前面提到的小偷案例具有同样的道理。如果小偷能按照我们上面的假设采取理智行为，那么，他一定能够意识到我可能优先选择向他开枪；同时，他也一定能够意识到，我可能认为他会优先选择向我开枪而首先向他开枪。但是，假设我们事先知道这两个基本（外在）“可能”将是什么行为，那么，我们就不需要采取过激行为。这两个“可能”中的任何一个都足以令双方或至少是一方先发制人，导致双方都选择进攻，从而任何更多担心都显得多此一举；或者二者都不足以

令我们中的一方为了自卫而首先开枪，我们知道这一点。因而除了这两个“可能”之外，我们没有必要担心其他因素。如果我们都意识到，对彼此首先开枪的担心并不足以令双方首先开枪，那么，我们就能意识到双方都没有必要担心对方的自卫举动，也不需要担心对方会担心，等等。[1]

但是，如果我是出于紧张走火而非故意开枪，则情形将大不一样。假设我的紧张程度取决于我害怕的程度，而我的害怕程度取决于我担心对方会首先开枪；假设对方这是同样情况，那么，当我想到对方会优先选择向我开枪时，我就会越发紧张，结果我越有可能向对方开枪，尽管这不是我的本意。对方看到我的紧张，他也会变得紧张；对方的紧张反之又令我更加紧张，我更有可能向对方开枪。对方看到我更加紧张，他自己也会更加紧张；他的紧张又反过来加剧了我的紧张程度，结果我愈加可能向对方开枪。现在，我们可以把一方的紧张看做是对方紧张的结果，而且得到了两个同时进行的微分方程。这两个微分方程产生了一个特殊现象，而这正是我们所要分析的。[2]

之所以会出现这种情况，主要原因是我们设计的模型没有涉及决策的标准问题，即没有涉及一个能够告诉我们双方战略选择的行为假设。相反，我们的“紧张模型”（nervousness model）表明，当人们担心自己可能遭受进攻时，通常以先发制人达到自卫。我们能否仅以这种方式，即通过研究选手的决策概率而非决策规则——并非在选手可以完全选择并制定自己战略的模型中——得到我在本章开始时描述的现象。

现在，这是否意味着决策果断的理性选手不能演绎我们的假设场景呢？鉴于一个理性人也可能出现紧张，我们能否预测到选手采取史无前

[1] 例如，双方能够进行沟通并不断地交换意见，他们就会达成非正式的共识，选择不开枪，那么也就不存在欺骗的动机——假设双方知道彼此的基本参数。

[2] 在这里描述的问题当中存在一个重要的非对称现象。我们允许双方在不应该的时候首先开枪，双方也知道这一规则——在“恐惧”状态中；但与此同时，我们不允许选手错失开枪的良机，双方同样知道这一规则。（当然，也可能存在小偷的弹药进水或忘记上膛的可能性。或许我仅知道这只是个意外，对方也可能知道我知道这一点，等等。）这一可能性很显然能够促使双方保持冷静，减少双方开枪的可能性以及意外走火或非理性开枪的可能性。

例的行为应对复杂变化的环境或新信息？在这种情况下，我们的理论将涉及心理因素而非智力因素。但是，我们能否假设一个理性博弈选手看到前面提到的小偷时，将调整其转轮的设置？[1]

当然，个体决策与集体决策在这方面存在很大差异。尽管不同的成员具有不同的价值观，对来自外界的突袭采取不同的反应，但是我们可以通过选举实现集体决策。由此可见，赞成进攻的投票数取决于未来受到攻击的几率。假设投票数还依赖于一些偶然因素，如：选举当日的缺席人数；倾向于进攻的有关当局日益受到对方发动进攻的概率的影响，反过来，前者的集体决策也逐渐增加对后者的影响。因此，如果我们假设存在一个具有选举机制和不同价值观的集体选手，那么，我们就能创造一个理性选手参与的博弈场景。

此外，我们还可以以另一种方式调整这一模型，使之适用于一个决策果断的理性选手。就伙伴问题和突袭问题而言，这一模型可能具有相当普遍的代表性，而且涉及了实际军事突袭问题最核心的部分，即决策对不完全预警系统的依赖以及在决策过程中出现“1类”和“2类”错误的概率。

不完全预警系统产生的随机行为

尽管预警体系的应用可以有效地降低遭受袭击的危险，但是预警系统也可能出现差错：预期警方系统可能使我们误把敌人的战机当做海鸥，从而错失防御良机；或者，预期警方系统可能使我们把海鸥当做敌人的战机，促使我们对敌人发动意外攻击。当然，如果我们肯投入巨额资金和精力用于改进预警系统，那么，我们或许能够避免出现上面两个失误。但是，就固定有限的预算资金而言，决策的标准往往是一方面的投入增加必须以牺牲另一方面的投入为基础，结果导致一个失误消失，另一个失误却愈演愈烈。在采取报复之前要求减少有关对方发动进攻的证据，无异于要求把对方的战机当做海鸥而撤回自己的战机。

[1] 请注意混合战略的一般规律——理性地调整某一方的决策转轮——与这里讨论的场景没有关系。

但是，我们现在为理性决策者建构了一个模型。在这个模型中，理性决策者不必做出明确的决策，采取某种具体行动或放弃反击以应对对手的潜在进攻，而是通过调整自己发动错误进攻的可能性事应对。提高预警系统的决策标准，降低决策失误的概率无疑有助于加强应对潜在进攻的防御能力，但是也提高了导致报复行动的假警报的概率。如果双方都以提高自己的意外进攻能力来应对日益严重的突袭，那么，一方发动进攻的可能性则来自于对方发动突袭的可能性。[1]这样一个预警系统无疑将是小偷场景的理性、机械翻版。

在这个对称性模型中（为了易于操作，我们建立的模型具有对称性），我们假设 h 代表赢得战争的收益，–h 表示输掉战争的损失，0 代表双方同时进攻时的预期收益（赢输概率均为 50%），1.0 表示维护和平的收益。[这里，我们允许 h 的值大于 1，只要（1–R）h 在收益矩阵中的值小于 1。但是，如果胜利者付出极大代价而获得胜利，那么 h 的值将相应减少。] 我们假设突袭成功的一方赢得最终胜利的可能性非常大——突袭成功意味着在对方没有进攻时，一方发动突然袭击；或者突袭没有被对方的预警系统发现。 R 代表双方预警系统的性能，即能够准确预期判断对方的进攻，并以先发制人的方法进行防御。收益矩阵如图 9-6 所示：

	I	II
i	0 (右上)；0 (左下)	–(1–Rc) h (右上)；(1–Rc) h (左下)
ii	(1–Rr) h (右上)；–(1–Rr) h (左下)	1 (右上)；1 (左下)

图 9-6

即使理性思维反对发动进攻，选手也可能发动袭击， 即存在在不应该的时刻发动袭击的可能性。这种可能性包括两个部分：一部分以 A

[1] 正如以后所涉及的，并不一定需要出现这样的现象。假使受到袭击的危险性不断增加与对方突袭能力的减弱有关，则一方完全可以按照本文描述的相反原则做出反应。

表示，非理性进攻来自外界因素的可能性，并排除了假警报引起进攻的可能性；假警报可能引起的进攻以 B 表示。因此预警系统可能导致的两个错误分别以 B 和（1-R）表示；模型的主要特点为 B=f（R），f′（R）>0。这样一来，我们在降低了（1-R）的错误概率的同时，反而提高了 B 的错误概率，反之亦然。

B 和 R 的值影响了双方的战略选择，因为前者能够将后者的损失最小化，即将其收益最大化。假设 Vr 代表 R 在游戏中的预期收益，R 的预警系统问题在于必须按照 B=f（R）选择 B 和 R 的值，最大值为：[1]

$$
\begin{aligned}
Vr=&(1-Pc)(1-Pr)+Pr(1-Pc)h(1-Pc)\\
&-Pc(1-Pr)h(1-Pr)\\
=&(1-Ac)(1-Bc)(1-Ar)(1-Br)\\
&+(Ar+Br-ArBr)(1-Ac)(1-Bc)h(1-Rc)\\
&-(Ac+Bc-AcBc)(1-Ar)(1-Br)h(1-Rr)
\end{aligned}
$$

除此之外，根据前面的矩阵分析， R 应该检验调整后的收益矩阵，以证实和平共处是否仍是双方的最佳选择，因为这一矩阵来源于 Rr 和 Br 的“最佳”值以及 Rc 和 Bc 的预期最佳值（观察到的值）。实现和平共处和改进预警系统的条件是：

$$Pc=(Ac+Bc-AcBc)<\frac{1-h(1-Rc)}{1-h(Rr-Rc)}$$

$$Pr=(Ar+Br-ArBr)<\frac{1-h(1-Rr)}{1-h(Rc-Rr)}$$

由于对称性的原因，右边的分母趋向 1。

实际上，正如下面所要看到的，第二次复审根本没有必要。在某些行为假设中，R 和 B 的“最佳”变化（取任何小于 R=1 的值）都要求必须满足保持修改后的矩阵稳定的条件。

我们需要进一步详细说明选手的实际行为。广而言之，针对“参变

[1] 之所以如此假设，是为了易于说明假警报导致的意外袭击与有预谋的人为进攻没有区别，同样能够达到突袭的效果。尽管我们忽略了可能被想当然地认为是单位时间内出现假警报的概率的时间因素 B，但是（1-R）依然可以表示进攻的错误概率， A 或许也存在相关因素。鉴于此，我们在这里假设了时间因素的存在。

行为”（parametric behavior）、“默式博弈”和“谈判博弈”的区别我们可以进行 3 个假设。

参变行为

首先，我们假设双方都将遭受对方进攻的可能性看做是一个参数，而非表示损失的变量，同样也将对方的预警系统看做是一个参数。这意味着一方可以直观地看到自己的 R 和 B 的值，并主动为 R 和 B 选择最佳值以最大限度减少自己可能遭受的损失。这一假设可能会使双方的选择 B 成为对方进攻几率的上升函数。（这一假设只存在“可能”，而非一定，因为另一个选择 R 也可能增加动机函数，从而抵消 B 的作用。下面将会涉及这一问题。）如果我们假设两个选手一直根据遭受袭击的参数不断地变化自己的 B 和 R 的值，并密切关注对方的 B 和 R 的值的变化，但是双方都没有将对方的行为参数作为自己的函数。那么，我们就得到了一个动态“乘数”系统——保持稳定还是变化剧增取决于参数的取值和？函数的状态。我们可以取双方的 B 的最佳值作为对方的函数，从而通过解下面的两个等式推导出平衡状态的稳定条件。我们同时计算与双方的 B 和 R 的变化有关的“乘数”在 f 函数的范围内移动或在 A 参数的范围内变换。

简单地说，为了使选手 R 找到一个适合的“参变行为”函数，我们使 Vr 随 Rr 而最大化，同时 $Br=f$（Rr），Bc 和 Rc 保持不变。Vr 的公式表达为：

$$f'=\frac{Pch(1-Br)}{(1-Pc)[1-h(1-Rc)]-Pch(1-Rr)}$$

同时，h（1-Rc）<1>h（1-Rr），$f''>0$。

由于我们假设 f′为正数，分母也必须为正数。如果 R<1，Vr 可以得到最大值；但是条件是 Pc 必须满足分母为正数的这一条件，只有这样 R 才会优先考虑和平共处。如果 R<1 时，双方都能做出最佳调整，调整后的 B 和 R 值必须满足和平共处的条件。

在这个行为假设中，Br 和 Bc 的关系——产生 R 的 B 最佳值的结果函数的斜率——可以通过调整上面等式两边的项得到：

$$\frac{DBr}{dBc}=\frac{dBr}{dRr}\frac{dRr}{df'}\frac{df'}{dBc}$$

$$=\frac{f'}{f''}\frac{df'}{dBc}=\frac{f'}{f''}\left(\frac{\partial f'}{\partial Rc}\perp\frac{\partial f'dRc}{dBc}\right)$$

$$-\frac{f'}{f''}\left(\frac{\partial f'}{\partial Bc}\perp\frac{\partial f'/\partial Rc}{\Phi}\right)$$

Bc=Φ（Rc）表示选手 C 的相应函数。

鉴于$\partial f'/\partial Rc$是负数，取值更小的 Φ′可能使选手 R 的 dBr/dBc 为负数。但是，提高意外袭击的代价足以抵消遭受进攻的可能性。换句话说，Br并非仅仅是 Bc 的函数，也是 Φ（Bc）的函数； Br 可能随着 Bc 的增大而增大，随着 Rc 的增大而减小，而且随着 Rc 和 Bc 的增大会逐渐脱离 Bc 轴。

一个稳定的平衡状态要求，选手 R 的 dBr/dBc 和选手 C 的 dBc/dBr 得到的值必须小于 1 即以垂直线上的 Br 和水平线上的 Bc 为两点，C 的抛物线应该与 R 的抛物线的下端部分相交。将 B 和 R 的变化与函数的变化联系起来（或与 A 的值的变化）的整体“乘数”表达式包括 1 减去分母中的结果。

正如前面提到的，表达式f'的分母消失了，结果 Rr 和 Br 以及f'的值迅速增大，h 的值逐渐满足矩阵不稳定的条件。（实际上，矩阵博弈的稳定与参变行为平衡状态的稳定不同。前者与参变行为假设存在联系；预期矩阵收益和预期判断对方行为则是为了说明对方的行为，而不是为了关注或想方设法适应后者。）

我们必须注意到选手 R 有可能在其考虑过程中忽视 Ar。Ar 跨出公式是为了使 Br 和 Rr 取得最佳值，因为 Rr 或 Br 的值能够出现差异的条件是选手 R 无法进行“非理性”袭击。如果 R 发起了非理性进攻， B 和 R 的值将对其没有影响。（然而， Ar 确实能对矩阵的稳定性产生影响，因为 Ar 成为 Pr 必须满足的条件。因此，为了说明 C 的变化，R 不得不考虑 Ar 的作用。但是，仅仅说明 C 的变化，而不连续关注 Bc 和 Rc，将使 R 的行为非参变化，这将与现有的假设相矛盾。如果 R 考虑增加投入用于改善自己的预警系统，那么 Ar 将对 R 的决策产生影响，

因为它将导致系统产生其他结果。这一考虑显然已经超出了现有的模型范畴。）

默式博弈

现在我们进行另一个行为假设，并将产生与前面同样的结论。这里，我们不假设双方能够看到对方如何变化 R 和 B 的值，并做出相当的选择。现在，我们假设双方都知道对方的技术机会（technological opportunities）——R 与 B 对另一方的函数关系——但是无法知道对方是如何变化 B 和 R 的值。这意味着双方都知道对方的预警系统的运行机制，但是无法确定对方的预警系统如何按照其指令处理传入的信息——即对方的决策规则。这个假设为我们提供了一个非合作博弈场景。一方必须在知道对方的收益矩阵，而不知道其选择值的情况下，确定 B 的值（或 R 的值）。

在这种情况下，我们得到一个具有典型“平衡点”的收益矩阵，即参变行为假设产生的一个稳定平衡状态。[1] 换言之，参变行为假设中的解决方案仍然可以被称为非合作博弈的解决方案。（在两种情况下，平衡点不具有惟一性。如果不是惟一的话，第一个假设使结果取决于初始条件和“突袭”；第二个假设则更趋向于将明确解决战略方案的智力问题复杂化。）

当然，这一解决方案无法满足选手双方的要求。这是我们前面提到的“囚徒困境”的典型。双方不断增加 B 的值无疑增加了互相进攻的可能性。[2] B 显然还存在一个对双方都有利的更小值。假设双方人为发动突袭的概率相同（双方 A 的值相同），那么，双方之间就消除预警系统（消除假报警的可能性）达成一个协议，无疑是二者的上乘

[1] 博弈论中平衡点是指选手双方做出的一对战略选择，一方的战略选择与另一方对应的战略选择搭配时才能发挥最大作用。（对方选手双方而言，或许存在数对平衡点。）

[2] 经济学家们或许从两个生产商那里发现了一个经济现象。这两个生产商合理分配双方有限的生产资源用于两种产品的生产。一种产品是涉及外在节约因素的“防假警报安全”，一种是涉及外部不经济的“防突袭安全”。

选择，如果他们受到限制必须拥有同样的预警系统的话。[1]

谈判博弈

如果我们考虑到双方能够以谈判消除预警系统的消极因素，为了共同利益牺牲较小的 R 来实现 B 值的减小，而且这样的协议也具有现实可行性，那么，我们完全可通过谈判框架具体化推导出一个惟一的解决方案。假设解决方案和博弈场景具有对称性，即双方就共同的问题——R 和 B 的值——进行谈判，谈判的结果是 B 和 R 均为 0，双方之间根本不会存在所谓的预警系统，因为如果双方的预警系统相同，那么双方进行突袭的初始可能性将存在重大差别（在 Ar 和 Ac 之间）。超过这一点，一方将为消除预警系统协议付出代价。

但是，一般而言，这会演变为一个内容广泛的谈判问题，甚至将比现在的公式涉及的内容更广泛，因为双方可能不仅仅要思考如何变化 R 和 B 的值，而且还可能以直接进攻威胁对方或者通过机制安排决定 A 的值。

尽管双方为了共同利益就降低 B 和 R 的值达成协议，但是在具体执行过程中会遇到很多困难，因为一方可能无法得知对方确定的 R 值和 B 值。他们需要——至少在非常重要的程度上——协调未来决策的标准，不单单是一个摸得着、看得见的预警系统。他们需要知道，双方在紧急事态中需要多长时间才能“确定”并能够承担多大程度的风险。假设模型代表的是一场全面战争而非边界蓝图或一方对另一方的小规模进犯，那么，一旦协议失败，并导致某种后果，如战争，那么反责和损坏

[1] 假设 A’s、B’s 和 R’s 的值相等，那么 Vr 和 Vc 等于（1-P）^{2}，其中最大值为 B 等于 0。（如果 B 的最小值大于 0，则取决于 A 的值。）如果 B’s 和 R’s 相等，但是不等于 A’s，那么

$$dV_r/dB=-2(1-B)(1-A_c)(1-A_r)+(A_c-A_r)(h/f')$$

可以为正数，其中 Ac 大于 Ar 和 f'。在这种情况下，选手一方——拥有小 A 的一方——更倾向于依赖某种预警系统，尽管与没有时相比，双方都倾向于使用预警系统。但是与参变行为相比，这使 B 和 R 的值相对减小。尤其是当上面公式的结果为 0 以及把 f' 公式与相应的参变行为的结果相比时，这一现象更加明显。

赔偿诉讼一定会在双方之间发生。

情况也可能是 R=B=0 本质上能够被觉察——不存在任何物理形式的系统。即使有效的强制性系统可能不存在，如果 R=0 时矩阵不稳定（当 h>1 时），那么，以 B 形式存在的一些风险因素足以将 R 置于安全范围，即 h（1–R）小于 1。

对双方而言，达成一个二者都承认 A 的协议也并非易事，因为承认对方的 A 大于 0 在政治层面很难实现。

鉴于此，选手双方不得不倾向于此前明显削弱二者突袭能力或者说改善双方的转换抛物线 R 为（1–B）的安排。例如，双方达成共识，投入更多资金用于改善预警系统的性能；或者实力相对较强的一方可能为对方提供人力和物力的帮助。这样就避免了一方对来自对方的袭击的担心，以及另一方遭受假警报的困扰。选手双方更有可能就优化双方的兵种结构达成某种协议，因为这样不仅可以消除潜在的突袭，而且提高了彼此军事力量的透明度，从而能够有效地防止来自对方的突然袭击。对双方来说，他们不需要将 R 和 B 作为协议的一部分；相反，由于 R 和 B 的隐蔽性影响了函数 f 和 ø，选手双方必须考虑这些函数都涉及各自的预警系统以及对方的进攻力量。（但是，我们必须注意到预警系统——在 R 值一定的条件下，函数 f 和 ø 的值随着 B 值的减小而波动，反之亦然——性能的“提升”只是在某些情况下才能保持平衡。 R 的边际成本的提高导致 B 值的增大，双方可能共同意识到提升预警系统性能的必要性，这如同“发展”了“囚徒困境”矩阵，即提高双方的非合作战略选择的收益。）

谈判博弈公式化将我们引入了对谈判策略的分析。例如，假设一方采取参变行为，对方知道后也将考虑这一因素，那么第一方会设置一个“反应函数”（reaction function）V 并引入到对方的公式中，成为对方试图最大化的函数。一般而言，本书第 2、5、7 章中所论述的诸如对“战略行为”的研究与突袭、伙伴原则博弈有着密切联系。

多人博弈

我们上面探讨的博弈场景大多涉及两名选手，即：二人博弈。但

是，当参加博弈的选手不断增加或存在第三方选手作为独立实体被引入博弈场景时，我们的研究对象将出现一个有趣的新变量。某种程度上，选手双方不仅需要考虑来自于对方的袭击，还必须考虑和预防来自第三者的进攻。这样一来，双方在二人博弈中消除预警系统的积极性和信心无疑受到了严重打击。毫无疑问，在一个人数更多、规模更大的博弈场景中，任何两个选手都会发现共同改进彼此的预警系统，降低警报失误的概率对自己来说都是利大于弊，因为二者均考虑到了他们采取参变行为时所忽略的“外部管理失当”（external diseconomies）因素。例如在下面的场景中，两名全副武装的保卫人员同时在同一栋建筑内巡逻，二人随时会向自己看到的可疑物体开枪，其中当然包括对方，那么二者如何才能避免误伤自己人呢？为了避免此类事情发生，双方最好达成某个强制性协议，规定双方必须弄清可疑物体后方可开枪，只有这样才能降低误伤发生的机率。（实际上，假设我们将前面模型中的 P_c 和 P_r 分别表示一人在黑暗中遇到小偷的概率，而非两个巡逻人员，那么，二人巡逻场景将是我们前面提到的模型在现实中的翻版。为了增加博弈场景的复杂性，我们必须为小偷的行为引入相关参数——让小偷作为一个理智的第三方出现在场景中，并努力预测双方的决策选择。[1]

[1] 澳大利亚国立大学的亚瑟·李·伯恩斯（Arthur Lee Burns）探讨了有关三人或多人博弈场景。当明显的歧义行为被引入到相互猜疑的行为模式中时，由于受到第三方的恶意挑拨，往往在其他两名选手之间发生人为战争。特别是当一方出于技术原因或共同监督，认为预警系统有利于一个或两个核心选手监督另外一方的雷达时，这一分析得到更有价值的结果。详情请参阅伯恩斯的“催化战原理”〔“Rationale of Catalytic War”，（Center of International Studies，Resarch Memorandum No.3;Princeton University，1959)〕。

第10章

突袭与裁军

尽管“裁军”通常都是敌对各方为了实现缓和，减少战争爆发的可能性或缩小战争规模和程度而采取的自卫行为，但是“裁军”方案往往各不相同。有些裁军方案是各方理智的抉择，而有些则是一时的感情用事。裁军方案大多都涉及对武器数量和质量的限制，特别是“进攻性武器”和一些可能有意或无意中引起民愤或造成公共损害的武器。有些裁军方案异常复杂，内容全面而广泛；而有些方案则仅仅涉及双方共同利益特别明显的某个领域。谈判各方往往只需要一点努力就可以在该领域达成共识，而且一旦成功则为全面裁军开创了良好开端。自从美国总统艾森豪威尔在 1955 年首次提出“开放天空”方案以来，谈判更多地涉及防御突袭的非全面方案或措施，并且愈加重要。

裁军各方对突袭的重视，并非意味着各方失去了对销毁大规模武器的兴趣，反而更加表明了各方面希望在一些容易取得共识的领域实现合作，从而建立成功合作的传统。不管是美国政府还是其他国家政府，都把共同关心的防御突袭问题看做是推动裁军向前发展的有效途径，而非寻求裁军的替代品。

此外，虽然传统的裁军方案往往涉及当事方如何预防突袭，但是其中仍然不乏创新的思想。即使是正统的“开放天空”方案也并非新瓶装旧酒。其根本思想依然认为，只要武器还处于备用状态就不应该具有挑

衅色彩——即这些武器仅仅是用于威慑而非进攻性。与传统观点不同的是，该方案并不主张必须适度地对敌人保密，使对方在某种问题上对我们的意图琢磨不透，存有猜疑；相反，该方案认为，更为重要的是如果我们实际上没有进攻对方的意图，那么，我们必须让对方知道这一事实，而不应该让对方存有疑惑。我们不仅要亲眼看到对方没有进攻我们的意图，而且希望对方也能亲眼看到我们没有进攻他们的企图。

在某种程度上，对对方采取必要的透明策略如同我们宣布自己缺乏进攻对方的政治支持一样重要。正如莱斯利·R.格罗夫斯将军（General Leslie R.Groves）在其演讲中所指出的："如果前苏联人知道我们不会首先对其发动进攻，那么，克里姆林宫（Kremlin）也不会进攻我们……尽管我们的消极防御对我们不利，但是这是目前阻止世界大战爆发的一个有效方式。"[1] 在我们生活的这个时代，任何一方都不甘心落后于对方以避免陷入不利的境地，结果任何微小的袭击行为都可能导致一场全面战争的爆发。假设我们不断担心，对方为了阻止我们对对方发动进攻而主动进攻我们，"自卫"（self-defence）将变得异常复杂。被看做是双方相互猜疑和交替恶化的"自卫"问题也向我们表明，我们不仅需要向对方公开我们不愿保守的秘密，而且还应该向对方透露自己希望拥有的军事力量。

当然，如果对方也能采取同样的做法，情况将会更好。因此，我们完全有理由将突袭问题看做是双方可以通过谈判解决的问题。

随着科学技术的突飞猛进，突袭的方式和手段也发生了巨大变化，取得了重大进步，尤其是突袭方案的改进和进攻武器的改良。当然，随着反突袭手段的进步，突袭行动也变得愈加困难，而且先发制人战略逐渐丧失了其自身优势。假设先发制人的优势丧失殆尽或受到严重削弱，则不存在所谓的进攻动机。

美国和前苏联各自拥有的军事力量足以摧毁对方；或者如果一方对对方发动大规模核攻击，那么对方将采取同样的回击，甚至更大规模的报复打击已经成为一个不争的事实。但是，如果一方能够将对方彻底消

[1] *The New York Times*, *December* 29, 1957, P.20.

灭，那么，谁首先发动攻击又有什么关系呢？毫无疑问，答案在于我们并非特别关注自己是否能比前苏联人多活一天。我们担心的是，一次突袭能否彻底地摧毁对方的报复力量，从而避免对方的大规模报复行动。但是，问题是我们的现有力量并不足以将前苏联彻底消灭，完全摧毁其报复力量，而只能保证我们遭受进攻后进行必要的还击报复。因此，我们必须假设如果前苏联先发制人而对我们发动进攻，那么，他们将首先打击我们赖以生存的报复力量。

这是两个完全不同的场景。一个场景是双方可以彻底消灭对方的“恐怖平衡”；另一个是不管谁先发动进攻，都可以将对方彻底摧毁。这并非构成相互威慑的“平衡”——双方条件的完全平等或对称；而是平衡的静态化（*stability*）。只有在无论哪一方先发制人都无法彻底摧毁对方的报复力量的情况下，这种平衡才能呈现静态化的趋势。

另外一种所向无敌的进攻性武器说明了静态平衡与动态平衡之间的区别。[1]西方旧社会的“平衡机制”（equalizer）使一个人杀死另一个人成为可能，但是没有想到可能出现双方同亡的情况。我们几乎每天晚上都可以在电视中看到这种武器系统的巨大威力。先发制人的优势鼓舞着每个人都采取先发制人战略。幸存者或许会为自己辩解说：“他可能出于自卫先杀我，所以，我只能出于自卫先杀他了。”或者“他认为我可以出于自卫首先杀他，所以，他想出于自卫首先杀我，因此我只能出于自卫首先杀他了。”但是，如果双方都留下一口气，能够以一支完好无损的胳膊回击对方，那么，先发制人也就失去了优势，双方都没有必要担心对方会率先开枪。突袭之所以存在，主要基于被突袭方存在报复行为的可能性。即使双方的武装力量都坚强无比，势不可挡——双方都自信自己的力量能够赢得战争，并幸存下来；但与此同时又无法彻底摧毁对方的报复力量——那么双方也不一定冒着风险，首先对对方发动突袭。这样的话，另一方也就没有必要过于担心突袭的降临，从而避免了

[1] 一位军事历史学家谈到“历史事实”时表示，人类历史上从未出现过一种攻无不克的武器。这无疑提醒我们“人类经过500年的火制武器后，还没有找到合适的方式对付子弹”。伯纳德·布罗迪，《绝对武器》（*The Absolute Weapon* [New York，1946，pp.30~31]）。

假警报的发生。

因此，针对突袭的防御计划关注的首要目标不是国民的安全，而是自身防御武器的安危。与其他裁军方案恰恰相反的是，突袭方案建立在威慑的基础之上，旨在防止对方发动袭击。突袭方案是为了完善和稳定相互间的威慑——提升某些武器系统的整体性能；突袭方案关注的往往是那些具有最大杀伤力，以及对公众生命财产构成最大威胁的武器系统——这些武器大多属于防御性武器，其使命主要是为了惩罚进攻者，即事后对进犯者予以沉重打击，而非事前迫使对方解除武装。从根本上讲，只能伤及民众而无法打击侵略力量的武器属于防守性武器。面对拥有类似武器系统的对手，一方显然不会先发制人而首先发动进攻。人们设计这类武器的初衷主要是将其用于打击军事目标—— 打击敌人的导弹或轰炸机——从而削弱对方先发制人的优势，动摇其先发制人的动机。

在确定了突袭问题的核心是双方的报复能力都可能遭受对方的攻击之后，现在我们可以明确反突袭措施与传统的裁军方案的最大不同之处。同时，我们还找到了大量异常情况和疑问的根源，这些都是我们认识某些方案本质和缺陷以及理解这些方案背后隐藏的动机的必要条件。完成上面的工作后，现在我们开始思考，反突袭方案能否被看做是实现传统意义上的全面裁军的“第一步”（first steps）；或者与其他裁军形式截然不同。保卫战略空军司令部（SAC）的措施能否被看做是实现武器销毁的“第一步”？为了实现相互威慑，我们能否从一开始就采取合作以维护和提高双方的大规模报复能力，并希望借此将大规模报复威胁从这个动荡不安的紧张世界消除掉。

或者我们能否将反突袭手段看做是双方的一种妥协——即默认“相互威慑”的存在是实现军事稳定的条件；这也是我们将要找到的——或者承认我们至少可以通过某些努力实现静态平衡取代动态平衡，尽管我们无法以更好的方式取代恐怖平衡。[1]

[1] 假设读者因为美国的报复力量足以抵抗任何外来突然袭击，而认为我这里提出的观点原则具有正确性，但是实际上却枯燥无味，那么，我希望这些读者进一步参阅 Albert Wohlstetter，“The Delicate Balance of Terror”，Foreign Affairs，37：211~234（January，1959）。

一旦确定了突袭问题的核心是双方的报复能力都可能遭受来自对方的攻击，在记住这一战略弱点（strategic vulnerability）的同时，我们有必要进一步评估军事力量、防御手段以及核查军控方案的价值。例如，我们不会以美国和前苏联各自拥有的轰炸机、导弹、潜艇和航空母舰的数量来评价双方的战略力量，这看起来似乎是我们想知道谁更能够在和平年代进行一场规模盛大的阅兵仪式。通常情况下，在军备竞赛中领先的一方总是先发制人，率先进攻。如果我们总是考虑对方会首先发动进攻，那么，我们就会认为即使是 2000 架轰炸机只剩下 10%，即 200 架完成了报复行动，这也是值得的。

如果将研究过分局限于威慑，那么，我们会发现评估防御手段的价值将十分困难。尽管我们无法将芝加哥整个城市掩藏某个防空洞内，或者将其从现在的位置上挪 10 英里，但是掩藏、分散、坚固的隔离屏障以及空中巡逻机都是保护威慑力量的有效手段。当然，无论芝加哥的空防工作多么完善，整个城市也仅有 50%的机会从敌人的数兆吨炸弹中幸存。尽管这听起来有些令人气馁，但是更严重的问题是我们现在还无法做到这一点。我们必须承认，积极防御能够保证大部分战略攻击力量免遭攻击，甚至足以对前苏联进行有效的大规模报复行动。同样，芝加哥的防御能力不足以要求前苏联必须以三倍的力量发动首轮进攻才有可能取得成功；其意义仅在于，前苏联必须投入更大的兵力而已。但是，我们的防御力量则要求前苏联必须以三倍的兵力发动进攻，而这无疑增加了其进攻力量穿透我们的预警系统的困难，更大的意义在于前苏联将遭受大规模军事报复。

我们可以以同样方式评估军控的价值。如果仅仅是为了分析前苏联人进攻美国的目标城市，那么，我们是否知道前苏联人发射其洲际弹道导弹（ICBM）的具体位置已经变得不重要了。因为与袭击一个城市的数兆吨炸弹相比，其精确性已经毫无意义。但是，如果对方试图摧毁埋入地下掩体内的一个导弹或轰炸机，那么，他们的武器必须具有高精确度。但是对于一个城市目标而言，袭击误差在 2~3 英里内完全可以忽略不计。尽管如此，如果是打击一个掩藏严密的报复性武器，对方仍必须同时发射数枚导弹才有可能击中目标。因此，在传统意义上，洲际弹道

导弹打击目标的地区性限制在裁军中并没有引起人们足够的重视。然而，就维持双方威慑力量的平衡而言——增强双方报复打击的能力——通过降低双方武器的精确度来区分二者的袭击目标，无疑对双方具有不同的意义。（对于没有掩蔽的飞机和导弹而言情况如此；不幸的是，城市目标更是如此。）

在有些问题上，过分强调突袭问题可能导致产生与传统裁军完全相反的答案，如双方可能接受对导弹数量的限制。（假设我们与前苏联达成了有关限制导弹数量的协议，并存在核查的可能性。）综合考虑人口目标和对方的袭击动机后，我们认为在对方结束第一轮进攻后，我们需要至少 100 枚导弹用于对侵犯者进行有效的报复打击——即削弱对方先发制人的打击优势。为了进一步说明问题，我们假设对方导弹的攻击精确度为 50%，换言之，对方需要 200 枚导弹才能将我们导弹保有量降低到 100 枚以下。假设我们拥有 400 枚导弹，则对方需要打击的导弹数量至少是我们保有量的 75%，即 300 枚；考虑到导弹本身 50% 的折损率，那么，对方需要发射的导弹数量至少是我们保有量的 2 倍，即超过 800 枚。如果我们拥有 800 枚导弹，那么，对方需要打击的导弹数量则至少为我们保有量的 87%，即 700 枚；考虑到导弹本身 50% 的攻击精确度，则对方需要发射的导弹数量至少为我们保有量的 3 倍，即 2400 枚。依此类推，防守方拥有导弹的初始数量越大，攻击方越需要发射数倍以上的导弹，才可能将对方的导弹保有量降低到安全数量以下。[1]

根据上面的结论，导弹数量的限制有利于维护双方的平衡稳定，而且限制的数量越多越好。迄今为止，有两个原因可以解释这个现象：首先，双方拥有的导弹数量越多，他们就会保留更多的导弹用于防御对方先发制人，并进行一定程度的报复。其次，双方拥有的导弹数量越多，他们反而会不断增加导弹的数量，试图在第一次攻击波结束后在数量上赢得对对方的相对优势。因此，随着双方拥有的导弹基数不断增大，他们以欺骗或掩藏或破坏协约等手段实现自己对对方的数量优势越来越困

[1] 这无异于意味着他必须一次性地将全部的导弹发向对方，或者持续不断地发射导弹，结果他将无法知道发射的这些导弹中，有多少击中并摧毁了目标。

难。事实上，如果双方的扩军计划超过了各自的预算能力；或者在预算允许的范围内，双方拥有的导弹数量已具相当规模，那么，由于经济条件的限制，平衡的稳定性取决于双方为了赢得主动而协调能够做到的事情和必须做的事情之间的关系。

这里，有个案例说明了“军备竞赛”与局势的动荡不安并没有必然的联系。如果双方拥有相同数量的导弹，随着双方不断发展新的导弹，一方彻底摧毁对方的导弹防御能力变得愈加困难。各自预警系统承担的责任也越来越大。如果双方拥有较少数量的导弹，二者拥有的导弹比例为 2:1 或 3:1，那么，拥有绝对数量的一方将赢得优势，并倾向于先发制人；相对弱势一方用于报复反击的导弹绝对数量则会相应减少。因此，双方完全没有必要为自己的相对劣势而过分担心；同样，也没有必要为自己的相对优势而沾沾自喜。

与其说是为了说明裁军问题，还不如说“导弹决斗”（missle duel）为军备竞赛提供了充分理由。但是，这一案例同时也表明了，根据静态威慑和反突袭方案的逻辑，增加军备与裁减军备之间的问题必须以个案分析。准确地说，裁军有助于稳定是一个史无前例的新结论。

我们对导弹潜艇以及相关反潜技术的态度，很大程度上取决于我们是否担心敌人发动进攻或突袭。如果多年以来，事实证明潜艇对反人类导弹具有强大的隐蔽性和攻击力，那么，我们将认为发展潜艇不是一件可怕的事情，而是一件令人欣慰的事情。假如实际上，我们最大的愿望是实现相互威慑以及维持双方之间静态的平衡，那么，我们希望双方都拥有相当规模的能够携带北极星导弹（polaris-type missile）并移动灵活的潜艇。如果事实已经证明了潜艇自身的隐蔽性和可靠性能，那么，它就不需要为了全面进攻而先发制人，也不用担心侵犯者会打击他们认为的防御力量。当然，我们如能摧毁敌人的导弹潜艇，而对方却没有同样的能力，或者这种力量的存在是我们无法改变的事实，那么，我们只能希望这种能够摧毁彼此的力量本身具有强大的抵抗力。根据这一观点，我们或许不应该期望自己独占“坚不可摧的”（invulnerable）核潜艇。实际上，假如我们既没有主观意图也没有政治动力进行首发一击，那么，对方如能确信这一事实无疑具有重大的积极意义。他们如能确信自

己能够抵抗我们的打击，这对我们来说也许并不是一件坏事，因为这阻止了他们出于遭受攻击的担心而先发制人。如果对方担心自己的战略力量将受到我们的攻击，那么，我们也将面临同样的问题。

这些思想也影响了我们对发展反潜技术的态度。海军迫切需要一种性能更高的反潜系统。毫无疑问，我们必须全力解决这一问题。与此同时，我们或许共同希望这个问题没有解决方法。如果情况果真如此（某种程度上的技术原因），而且潜艇在未来的10年内仍能维持较高的安全性能，那么，实现双方间的平衡稳定在技术层面上依然可行。假如事实证实潜艇失去了自身优势，那么，军备技术的稳定性要比我们的期望糟糕许多。我们必须设法（try）研发出一种新型反潜技术，因为我们担负不起前苏联人拥有一种新式技术，而我们对此一无所知；我们必须了解一些相关的反潜技术以保证自己的潜艇具有反潜能力。正如一个人可以与另外一个自己无法信任的人达成协议，我们可以寻求一个挡箭牌，因为我们知道自己的伙伴也在进行同样的努力，尽管我们希望不存在任何飞箭。[1]

既然我们已经将问题分析到了这种地步，所以，我们必须对该问题进行进一步深入分析。假设我们的问题是为了向敌人证明，如果对方首先发动进攻，我们将予以报复性打击——也是为了使对方相信，我们知道其知道我们相信自己的威慑能力，并拥有先发制人的能力——我们应该看到技术进步的积极因素，尽管这些技术进步有助于提高我们的报复武器反人类的能力。如果采取措施保证我们的报复武器在敌人的第一次进攻中幸存完全符合逻辑推理，那么，我们就有理由提高那些能够幸存的武器装备的性能。正如伯纳德·布罗迪（Bernard Brodie）所说："如果我们考虑威慑的特殊需要，即强调报复的惩罚性，我们就需要一种超脏弹（super-dirty bombs）。鉴于即使少量的超脏弹用于打击报复也会令对方胆战心惊，所以，双方都希望拥有这种武器，最好是在袭击出现之前，而且威力越大越好。"[2]

[1] 鉴于我们这里关注的是相关的原则问题，而非潜艇本身，因此我可以把无法提前在公海上侦察到敌方的信息看做是无懈可击的一种表现。

[2] 参见 Bernard Brodie, Strategy in the Missile Age（Princeton, 1959）, p.295。

当然，如果我们承认“恐怖平衡”（如果保持稳定的话）仅仅是一个古代机制——交换人质——的大规模现代版，上述推理的新颖性则会消失殆尽。在古代，人们为了表明自己履行诺言的决心，往往将人质交给自己不信任的“伙伴”；在现代，尽管敌对两国可能距之千里，但是军事技术的进步足以使双方将彼此国民的性命控制在自己手中。只要一方拥有足够的军事力量摧毁侵犯国及其人口以回敬对方的进攻，那么“恐怖平衡”将表现为双方默认将一切可能的人质进行交换。当然，我们也不希望通过交换如此多的人质以实现与某个对手达成某种共识。但是在一个自然状态的世界中，即使一方违背承诺造成损害，受害人也往往无处倾诉。在这种情况下，交换人质或许是相互质疑或敌对的两个伙伴实现合作的惟一方式。[1]

当然，我们上面所进行的这些探讨并非简单地为军备竞赛寻求合法依据。至少在字面意义上，这些推论表明全面大规模的“裁军”——甚至仅是选择性地裁减可怕的大规模性武器——只能造成局势动荡不安，而非长久的和平。要想通过裁军避免灾难性后果，则必须实现成功的彻底（*completely*）的裁军。尽管如此，裁军仍将是军控研究的重要领域，虽然军控与前面的论述水火不容已被验证。

有人认为，有必要将适用于先发进攻的武器与适用于防御报复反击的武器划清界线。一端是“完全”防御报复性武器：携带超脏弹的相对非精确性装备可以摧毁敌国的一切，保存完好或极其掩蔽的报复性武器除外。而且这些武器自身也被保存完好或掩藏得极其隐蔽，以至于敌人的打击力量也找不到。理想的状态是，这类武器应该在防御性报复打击中具有优势，而在进攻性打击中处于劣势；与此相反的一端是完全的进攻性武器：这些武器经不住二次打击，只能用于在对方的防御力量运行之前发现和攻击对方，因为只有这样才能发挥其自身优势。这些“先发制人”（strike-first）的进攻性武器不仅使其拥有者产生了强烈的进攻

[1] 这里有必要强调的是，我探讨的主要是重大突袭问题。“人质”观念说明民防政策取决于其与意外事件的关系程度——例如有限战争、第三方的恶意行为、次大规模报复行为等等。本章的最后部分涉及了突袭和其他军事意外事件的相关关系。

欲望，使其在局势不明的情况下选择先发制人，而非坐以待毙；并且这无疑是向对手表明前者意欲先发制人。结果，这些武器导致敌对国在此之前不敢轻举妄动，而且当对方认为这是我们的快速反应时，他们也会采取快速反应。

在“完全”进攻性武器和“完全”防御性武器两个极端之间存在大量进攻性和防御性兼而有之的武器装备。这些武器如能幸存也可作为防御性武器；但是如果用于进攻，则可给对方的防御性武器以致命一击。假设真采取严格的保护措施，大部分都属于这一类。因此我们不能简单地将进攻武器和防御武器划清界线，在研究突袭问题时厚此薄彼，区别对待。如果我们考虑消除所有打击对方防御力量的进攻性武器，或者陶醉于先发制人的快感，那么，可能就没有防御力量的生存空间。[1]但是，突袭谈判或许能够关注另一个极端。

适于向对方暴露的目标武器首先是进攻性武器。向前苏联人建议保护好其战略力量或者前苏联人建议我们保护自己的战略力量看起来似乎都是不正常的现象。通常人们往往主张放弃那些易于暴露的战略武器。请注意这与“禁止炸弹”在本质上有什么不同。无论言论对此如何宣传，至少这看起来主张威慑的力量应该得到改善，而非销毁。

其次，我们或许能够限制对方的防御力量而非其反人类能力的运用。除非双方能够达成共识，承认突袭方案旨在增强（而非削弱）双方的防御打击能力，否则根本无法实现这一目标。有关导弹性能和优势的讨论表明了这类限制条件的存在。

第三，双方可以探索某种合作方式或行动上的互谅以降低误解而引发战争的概率。假设美国人和前苏联人都知道事实的真相，而且双方都能单方面选择令人信服的行为模式，那么即使是信息的自愿交流对双方也大有裨益。这也是为什么提出在北极区域进行空中交通管制的重要原因，同样地，也存在对双方都有利的其他类似的行为方式。上述措施的积极作用——正如双方坦诚讨论进攻武器的消极因素——在于它们体现了一些通常只有在正式协约中才能看到的互信和互谅精神，并有助于推

需要说明的是，这里我们关心的主要是突袭问题。

动单方面的积极努力。

第四，某种危机或紧急事态的处理不当也可能导致意外战争的爆发。本章的后半部分将以大篇幅对此问题进行深入探讨。

第五，一定存在某种方式以减少突袭事件的发生，并削弱先发制人的优势。这使我们想起了前面提到的开放天空方案。

在过去的几年里，有关突袭问题的公共讨论大多涉及如何降低突袭的几率，而非如何限制突袭行动使用的武器。开放天空方案的要旨在于，对军事力量的密切关注有助于阻止双方发动突袭，即使处于战略劣势的一方也不能受到威慑。

自开放天空方案第一次提出以来，建构能够防止双方发动突袭的武器核查机制变得愈加困难。随着氢弹的发展，突袭需要的相关飞行器逐渐减少；导弹技术的进步缩短了导弹投入进攻与击中目标之间的时间；越来越多的导弹潜艇等机动武器系统受到对方的监控；这些军事技术的进步无疑增加了无条件全面核查实施的困难度，并不断削弱核查的实际效果。主张对军事力量的战略转移和重点区域进行监控，看起来似乎也已过时。现在的问题是大型国际组织对某国战略力量的集中核查可能披露出在数小时内出现可疑行为，或在数分钟内出现假警报。目前，还没有可靠迹象证实其可操作性。

这并不意味着核查计划没有成功的可能性，而只能说明核查方案的前景比较暗淡。但是，假设一方无法派遣侦察机24小时不间断地跟踪对方的战机、导弹和潜艇，无论它们前往什么地方，则另一方也可以考虑派遣飞机、导弹和潜艇前往最容易被监控的地点。如果限制武器的应用有利于核查任务的操作，那么核查或许能够获得一些有利成果。虽然将武器核查问题与军控问题结合起来的主张有实现的可能性，但是在具体操作过程中会遇到许多严重问题。

其中，最严重的一个问题或许就是如何处理武器核查与隐匿军事实力的矛盾关系。当导弹的攻击性能被不断改进，具有足够高的精确度时，仅仅依靠水泥掩体是无法有效保护自己的战略防御力量的；即使可能的话，那也需要投入巨大的成本。机动性和隐蔽性是防御力量自身安全的根本保证。假使敌人能够打击所有的固定设施，并摧毁一切能够击

中的目标，那么，被袭击方必须设法提高自己的机动性能。在某种程度上，对方可以通过一切监控手段获得我们防御力量的具体位置。

对双方军事进攻力量的核查计划，也可能以其他方式泄露被核查方的军事部署，从而使其处于不利局面。例如，众所周知，有时台风也可能阻止敌人的轰炸机进攻我们，这无疑是我们报复打击威胁的重要组成部分。如果遇到这样的情况，敌人发动袭击将是非常困难的，这取决于对方是否知道类似事件能否发生在我们身上，或者能够准确无误地预测自己能否借到东风。假如双方的战略部署遭受非常严重的传染病袭击，那么二者之间就可能出现局势动荡，而且双方都有可能将自己暴露于对方的眼皮底下。幸运的是——假设双方偶尔也会因为不可抗力而置于非警戒状态——双方都无法知道对方的偶然性非警戒状态。

最后，对双方而言，或许存在众多提前获得对方进攻情报的方法，这些预警系统的价值在于我们得到对方的进攻情报后能否采取有效的防御措施。我们或许派遣我们的预警部队率先进入战斗状态，但是如果预警系统提供的情报模糊不清，这一行为就失去了积极意义。结果，一个假情报导致了一场战争的爆发；而一个准确情报排除了最后一分钟的威慑。

另一种极端情况是，我们可以坐以待毙，“守株待兔”。如果我们的有效防御能够打击对方的进攻气焰——如果我们可以给予对方沉重报复的话——我们或许为了更好地阻止对方做出最后的选择，迫不及待地想告诉对方我们已经准备就绪。

一个非常重要的问题是我们如何才能进行有效的防御。如果答案只是简单的“保持更加警惕”，那么，我们为什么不能首先保持警惕？大多数情况下，一旦知道进攻的警报，人们往往会时刻保持警惕，仿佛进攻永远存在。或许我们的战略空军司令官长期以来一直致力于在进攻警报到来时，减少飞机的准备时间，提高飞机的起飞速度；或者关紧大门，将飞机严密地掩蔽起来；或者设法提高飞机的作战性能。总之，在突袭到来的短暂时间内，他们可以选择的余地并不大。

尽管如此，在突袭到来时，一个国家也可以进行某些暂时的防御准备。例如它可以深挖洞，广积粮，但是这并不是长久之计；它也可以从

战略防御力量安全地飞离地面，躲开对方的炸弹袭击，但是这些武器不可能长久地待在空中；它也可以让部队进入战备状态， 24小时警戒，但是总不能长期处于警戒状态；它也可以停止商务飞机运行以提高预警系统的可靠性，从而避免预警系统将自己的民用飞机误为敌机，但是必须为此付出高昂的经济代价。换言之，一方可以为了突袭进行暂时的准备，但是无法长期实践。

但是，还有一个问题尚待解决。即使采取上面提到的行为，我们又能够坚持多久呢？设想如果我们实际上并不能使所有的飞机而只能使一半数量的飞机处于空中，那么，各方面的代价（机械事故、燃料和人员）也过于高昂。尽管如此，当突袭来临时，我们仍然可以使大部分飞机升空。这也意味着，我们平时的防御准确无法令对方感到震慑，而我们面对突袭警报时的临时防御却可以使敌人三思而后行。这是否意味着对方看到我们的临时防御准确后会马上放弃进攻呢？或者对方以逸待劳，坐待我们的飞机燃料烧尽，飞行员疲倦了，飞机迫降后再行进攻？如果情况如此，我们为什么不在得到突袭警报时先发制人，首先向对方发起进攻呢？

“疲劳”问题困扰着每一个保持高度警惕的国家。有两种方式可以解决这一问题：首先，一方必须尽量优化自己高度警惕的工作，在保证持续有效的同时，避免出现“疲劳”，因为只有这样才能达到最佳的防御效果。其次，一方可以在防御力量足以遏制对方的进攻时，与对方进行必要的裁军谈判，这也是最切中要害的方式。假设我们的高度警惕工作只能维持数天，那么我们就有几天时间要求前苏联人裁减军备或进行有关的谈判工作。裁判的标准是既能满足对方的要求，同时也能保证我们的安全，保证回到“正常”规模，从而避免双方发生一场全面战争。这意味着必须重新设计和建构更大规模的反突袭方案，而不能仅仅满足于以前的政治性方案。这同时也意味着谈判并非仅仅出于对突袭的长期担心，而是在告诉对方，如果在规定的期限内不能达成有效的协议，从而避免突袭，那么，殊死一战将不可避免。

上面的探讨并不是说，过分的警惕是多此一举或者劳民伤财。这些讨论只是为了说明仅仅依靠预警本身还远远不够。过分警惕只是为我们

提供了一个机遇，能否抓住机遇还需要策略。为了应对紧急事态发生，需要提前做好充分准备。如果等到我们知道前苏联人马上要对我们发动进攻，那么，我们甚至没有足够的时间向对方发出最后通牒。在如此短的时间内，决定一个既能满足我们要求，也能被对方所接受的通牒不仅在人力上无法做到，在技术操作上也不现实，而且还需要相关的步骤才能达到目的。因此，只有我们事先做好了充分准备，才有可能向对方发出强有力的最后通牒。

我们通常以两个标准建构一个核查机制，或者评价一个核查机制的有效性。第一个标准是，这个核查机制能否突破核查对象的重重障碍，有效地查找事实真相；第二个标准是，当核查机制对核查对象有利，则核查机制是否有助于核查对象主动提供可靠的事实资料。上面的两个标准如同一个设法寻找罪犯的犯罪证据；另一个设法让无辜的嫌疑人为自己辩护。大约说来，一个是在缺乏确凿的有利证据的情况下，对无辜嫌疑人进行刑罪假设；另一个依赖于有利的确凿证据，而且与特定的情况密切联系，即揭露事实真相对嫌疑人有利。

两种情况的区别还与两类方案密切联系。一类方案是为了减少对人为袭击的担心；另一类是为了减少对意外战争或无意战争的担心。这些战争大多来自于情报失误，或误解对方对某些机械事故或第三者不幸遭遇或者其他情况的反应，如双方可能误解对方会发动两败俱伤的战争。如果是一场预谋已久的人为突袭，那么，战争的发起者会为自己的行为寻找一万个正当理由，并掩藏自己的真实意图。但是，如果是“意外战争”，双方都会想方设法将事实的真相以令人信服的方式及时传递给对方，以避免对方做出错误的抉择。

误解侵袭

大家不妨考虑下面这个问题：假设事实上我们没有侵略前苏联的意图，但是前苏联人却想当然地认为我们会进攻他们，那么，我们怎么做才能让对方相信我们不会进攻他们呢？换言之，假设前苏联人知道我们一直担心他们会发动对我们的袭击，尽管事实上他们没有这样的企图，那么，他们如何才能让我们相信他们不会进攻我们？

很显然，双方仅仅将事实的真相告诉对方是不够的。当然，在某些情况下，言语的沟通也能消除双方之间的疑虑。假设前苏联人——这是一个大胆的设想——的某个军事基地遭受到了核武器的袭击，但是前苏联人只是简单地安慰我们，那只是一起意外事故，并非来自于我们的袭击的先兆，那么，类似的做法无疑具有积极意义。然而，在大多数情况下，人们往往认为这不足以说明对方没有进攻的企图或者说对方没有采取敌对行为。尽管如此，还是有什么方式能澄清某些误解，如对彼此某些军事部署的猜测等。我们不仅要向对方证明自己没有任何进攻意图，而且还要证明我们实际立场是欺骗对方，如果他们相信我们并管制其武力。

有限战争中的误解

特别是在有限战争中，双方的某些行为可能被对方误认为是战略进攻。例如，假设我们派遣可以用于攻击前苏联军事基地的飞机沿着前苏联方向行进，那么，这可以被前苏联人误认为进攻他们——或者这些飞机从北美基地或中东的航空母舰上起飞飞往前苏联南部的邻国。同样，根据掌握的有限情报，我们也有可能将前苏联飞机执行一次有限战争的任务看做是对我们的海外基地和航空母舰的全面进攻。尽管如此，这只是一次有限规模的袭击行动，并非打击美国防御力量的全面进攻。

问题在于我们能否采取积极措施降低类似误解发生的概率，从而避免双方因误解而加快预先报复打击；或者因过分警惕导致假警报的出现。一方可以尽最大的力量向对方证明亡羊补牢——一旦全面战争爆发，部署在世界其他地方的军事力量将进行全面反击——为时已晚。

互存误解

正如格罗米科（Gromyko）在一次新闻发布会上所描述的另一番场景：

毕竟，所有流星和电子干扰都能反映在前苏联人的雷达显示屏上。假设在这种情况下，前苏联人的飞机载着核弹和氢弹飞向美国或者美国的其他军事基地，双方的空军部队都在北极的某个地区发现了对方，那么，我们将不得不面临这么一个现实：敌人正在发动一场战争，人类也将卷入核战争的漩涡中。

假设我们现在就面临上面描述的场景，那么，双方如何才能延缓或避免彼此之间的误解？如果双方都有重新以合理、令人信服的方式做出选择的机会，二者就可能达成共识，共同做出妥协让步。

谈判的环境或许并不令人顺意。最好的结果是，双方可能拥有数小时进行谈判；而最坏的结果则可能根本就没有任何谈判时间。谈判成功的条件分为两个部分：首先，一定要存在某种“解决方案”——某种行为方式能够防止双方趋向战争，通过动态的平衡让步缓和双方的高度警戒状态。在这一过程中，不能使双方中的任何一方赢得不利优势（dangerous advantage），而且必须将局势控制在双方的军事能力范围内；其次，双方必须能够以某种共同方式监督或监控彼此对诺言的履行情况。除非我们拥有可靠的手段监督对方的履行情况，否则我们不会履行自己的承诺。反之亦然。或许我们也希望采取某种欺骗手段，逃避对方的监控，但是在很大程度上，我们还是希望能够在这种情景下建立一个高效的监督机制，对双方都有强制约束力。只有这样，当我们履行自己的诺言时，对方才不会存在任何疑虑。问题的核心是如何保证双方履行自己的诺言。如果情况如此，那么，双方都会迫不及待地将自己对承诺的履行情况告之对方，如果双方都实际履行诺言的话。

鉴于允许核查人员工作的时间如此之短，这一例子充分说明了预先核查和验证工作的重要性；同时也表明双方必须事先进行有关的准备工作，进行详细的规划，优化飞行路线。只有这样，双方才能够以最有利的方式向对方提供相关的真实资料，如果需要的话。

与此同时，这一案例也说明了检验核查机制有效性的两个标准之间的区别。当然，研发一种万能雷达随时随地追踪对方——对方也能够以同样的方式对待我们——以避免偷袭并非易事。问题在于当双方都希望以某种可靠的方式受到一定程度的监督时，如何设计这种雷达系统。在某种情况下，我们也可以巧妙地躲避对方的雷达监控；而在另外一些情况下，我们或许故意派遣战机暴露于对方的雷达，或者采取某些长期方式，只要对方以同样的方式对待我们。

长期监控

危机和紧急事态与长期军控核查的区别在于，二者要求的当事方提供的证据以及其背后的行为动机不同。愈加“悠闲”（leisurely）的核查过程很大程度上取决于消极证据（*negative evidence*），即证据的匮乏。双方可以通过提高和加强系统的性能来减少这类证据事实；双方或许意识到，长时间地进行某种隐匿行为则可能增加逃避监控的困难。但是危机期间，一方会要求更多的确凿证据；而另一方则可能没有时间保持相对优势和及时更新系统。一旦系统无法正常运行时，一方不可能迅速检验系统性能或提高和加强其性能。结果，危机协议不得不依赖于积极证据（*positive evidence*）。一方不会要求对方提供有关的消极资料，而要求提供积极证据，即对方履行诺言的情况。危机时刻之所以能够出现这些证据取决于双方的动机——紧急形势对双方达成协议或共识的紧迫性——在危机时刻得到强化。

超前机制

为了防患于未然，至少为危机或意外情况做好准备，建构一个方便对敌双方进行沟通或监控彼此的灵活机制不失为良策。特别有人主张在双方同意的情况下，建构一个超前机制，或者增加额外的设施和核查人员，这无疑有助于核查机制在危机到来时发挥有效作用。换句话说，我们不应该仅仅以双方在“正常”状态时的动机来评价核查机制的作用。我们必须认识到，当形势迫使双方就军控或至少是暂时的军控进行谈判时，允许双方建构核查机制和沟通的时间并不多。

具体地说，当需要建构一个核查机制监督协议的履行以延缓核试验时，我们必须认真考虑双方在军事危机尖锐时，如何充分发挥核查人员和设施的作用。核查人员、人员的工作地点、通讯设备、技术培训和监控设备、工作人员的可靠性以及数量都不应仅仅以核实监控的标准来评价和建构；我们还应该考虑到，当双方都受到意外战争的威胁时，这些因素对核查方式、核实和通讯的不同作用。

前面的讨论表明，恐怖平衡的稳定性——人为偷袭的减少和假报警的避免——很大程度上依赖于我们能否与前苏联人共同实现某种军事机制。时间终将揭开双方的科技秘密，我们也会发现双方都拥有足够的防御力量以抵抗对方的进攻，不管对方采取何种行为，从而以有效方式实现双方间高度稳定的威慑平衡。同样，时间也可能把不利的秘密告诉我们，这也将敦促双方更加努力以更快的步伐寻求削弱防御力量的新途径，并致力于保持它们。当然，这只是一个希望——不是假设——即使我们与前苏联人以最大的诚意和高超的外交技巧实现共同合作和局势稳定。因此，我们可以说，合作可能促进稳定或可能无法实现。尽管如此，我们相信与前苏联人的某些合作或相互制约，无论是正式还是非正式，默式还是显式，都将对恐怖平衡的稳定性产生重大影响。其中存在的风险也非常之大。因而，虽然我们无法肯定双方提高彼此防御能力的积极合作一定能发挥积极作用，但是我们必须考虑这一因素，并质问我们自己是否真正需要威慑平衡的绝对稳定，如果我们能够做出选择的话。假设我们能够制定一个方案，并知道前苏联人肯定会接受这一方案，那么，我们是否真想实现一个长久、有效的反突袭方案呢？

尽管前苏联人不会人为发动偷袭，而且他们肯定我们不会先发制人后也放弃先发制人，无疑令我们备感欣慰，但是这意味着我们对大规模进攻的防御能力至少是建立在前苏联人坚信我们不会先发制人这一事实基础之上。假设前苏联人的防御能力足以抵抗我们的第一次进攻，那么，他们或许就不会相信这一事实。或许我们可以这么判断：除非面临严重的挑衅行为，否则我们不会在没有绝对把握摧毁前苏联防御力量的情况下，率先发动对前苏联的进攻。鉴于此，强大的战略空军部队或许也失去了其应有作用，或许这就是完全两极世界的最佳状态，这也是我们在现实世界中——现实世界存在一个“第三世界”，这为我们以有效威慑而非两败俱伤的突袭来遏制前苏联提供了广泛的舞台——无法负担的奢侈品。

假设前苏联人能够以他们喜欢的方式任意发动对我们的袭击，那么，我们能否威胁进行报复，而非区域性的消极抵抗？如果双方的战略

力量足以保护各自免遭对方的进攻，那么，除了促使双方长期对峙，保证双方共存外，这些战略力量是否还有其他作用？

除了上述的作用外，战略力量确实还能发挥另外一个作用，即保证双方能够对彼此发挥惩罚性的报复打击。假如造成前苏联和中国的城市瘫痪的主要原因是突袭造成的市民恐慌、经济损失、管理失控和民族耻辱，而非首当其冲的军事力量，那么，威胁仍然发挥了作用，尽管对方的战略力量十分强大。

假设“大规模”可以被释义为无节制报复，那么，随着我们失去成功摧毁对方防御力量的希望，大规模打击报复也失去了其自身的可信度。但是，如果我们将有限或渐进的打击报复看做是迫使前苏联人不要采取过分行为的手段，或者将有限战争扩大到前苏联境内并通过更大的威胁利诱使之成为区域性军事行动，那么，随着双方战略力量的增强，类似的报复行为和威慑将更加有效。同样的道理，所有可能演化为全面突袭的有限战争也将消失殆尽。次大规模的打击报复的风险也将比现在小得多，因为双方对全面袭击的恐惧也相应地降低了；也就不会有更多的人担心我们的有限打击可能被看做是全面战争的第一步。那么，前苏联人也不得不相信我们将有限打击报复误认为是同归于尽的序曲。

尽管进行了如此多的论述，我们并不认为，可能导致有限反打击报复行为的有限报复行为不一定导致大规模破坏行动，也不可能演化为规模越来越大的报复行动或者在有限范围内具有某种破坏性。有限报复战争或许并不比区域性有限战争容易，甚至可能更复杂困难。这里，我们并不认为，惩罚性的相互打击报复比区域性有限战争更能带给我们安全。我们只是认为在任何我们可以依赖于报复威慑的情况下，相互打击报复与全面战争相比则是更为可信的威慑行为。

只有当双方都存在进攻的可能性时，战略力量才能“中立化”并且具有惩罚功能，从而为威慑威胁提供支持。然而，随着双方的报复力量的不断加强，全面报复的可信度日益降低；与此形成鲜明对比的是，有限报复的可信度与日俱增。鉴于有必要进行第三区域的威慑，因此无论产生什么后果，我们都不能蔑视一个拥有战略空军部队的世界。事实证明，特定的威慑威胁（大规模的）比其他威慑行为（有限规

模的）更有效。

只有极度乐观的人才会相信，我们可以明确地接受或者拒绝一个完全持续增强双方报复力量的计划。第三区域威慑将产生什么后果和有限报复的可能性，则取决于我们对未来的期望。[1]

[1] 更多研究请参阅 Thomas C.Schelling and Morton H.Halperin, *Strategy and Arms Control*, The Twentieth Century Fund（New York，1961）。

附录

附录 A　核武器与有限战争
附录 B　博弈论的非对称性
附录 C　重释“非合作”博弈的决策理念

附录 A

核武器与有限战争

随着更小规模和更小杀伤性的核武器不断涌现，并被装备适中的地面部队应用到有限战争中，以及深水炸弹和适用于空战的核导弹的发展，衡量核武器的传统标准已经失去了现实意义，例如在有限战争中将核武器与常规武器区别对待。当然，也有人认为在有限战争中使用核武器将使我们在政治上置于不利地位，特别是当我们首先使用核武器时。甚至那些认为核武器造成的人员伤亡如同凝固汽油弹的作用一样具有道德优势的人，也不得不考虑到全世界反对使用核武器这一政治现实。

附录 A 将主要讨论如何区别核武器和其他武器，因为这涉及到我们在有限战争过程中如何处理同敌人的关系。为了实现战争的有限化或者更好地理解有限战争的含义，我们有必要认清核武器与其他武器之间的界限，尽管这并非现实界限而只是心理感知、墨守法规的暗示性界限。有人仅仅基于对武器性能的分析而非有限过程——即寻求有限战争的有限条件、影响这些条件稳定的因素、这些条件权威性的根源以及选手达成这些条件共识的环境和行为模式等——的分析，想当然地认为这些高度精确的小型核武器是火炮武器的一种，因为完全可以被应用到有限战争中。认为小型核武器“仅仅是另类武器”的观点依据是，假设衡量区别核武器和其他武器的标准中没有武器性能一说，那么，就根本不存在支持上述有限过程的依据。

现在，我们能否以同样的标准区别武器的合作者呢？前苏联人与中国人之间的差别如同核武器与常规武器之间的差别。同样的情况还有中国人和朝鲜人的区别；美国人和台湾人的区别；英国人和约旦人的区别以及埃及人和阿尔及利亚人的区别。除此之外，前苏联和伊朗间的数百公里边界线和其以南数百公里的边界线，以及鸭绿江上下游的边界线或者希腊和南斯拉夫两端并没有很大的区别的边界线。尽管如此，除了偶尔与此类似的横渡大河或测量高山的实际困难外，这些边界线在有限过程中还是发挥着巨大作用。

或许有人会说，这些都是“合法”标准，而且都是现实存在的，而区别核武器和常规武器的标准却是抽象的。但是问题在于，这些标准实际中并非真正合法，它们仅仅具有“合法性”，现实中并没有强制法律机制强制有限战争方承认这些政治“边界”或者国籍。前苏联人绝不会依法将进入其国界线看做是战争中质的变化——虽然是一个强烈行为，但没有进入其国界；假设我们故意跨过鸭绿江，那么，中国人绝不会进行合法报复（不仅仅是抵抗），他们绝不会放弃任何阻止侵犯的合法权力，偶尔允许别人通过；如果前苏联飞行员参加了针对我们的有限战争，那么，我们也绝不会依法承认这些前苏联飞行员，或者将近东进攻我们的地面部队认为是“自愿者”；任何禁止私闯国界的行为或者一个新国家卷入一场冲突的战略[1]都意味着核武器的引入。敌人反应行为最重要的一个决定因素是当敌人无法针对我们的间断行为进行有效反应或者做出增量反应时，他如何释义自己默许的行为。

假设在前苏联或中国附近发生战争，那么，导致前苏联或中国边界成为分界线的重要因素通常是没有其他可行性的边界线。西方国家的军队跨越前苏联边界无疑是挑战——主要是心理上的挑衅——前苏联领土主权的完整性，以及表明或者至少暗示了进一步前进的意图，除非能够发现边界内前苏联人能够容忍外人跨越的“显著”界线。这样的话，前苏联人就知道我们跨越其边界的某个位置时表示我们没有扩大前进的企图；我们也能够清楚，假使我们跨越了其边界，前苏联人能够容忍的程度。而且，双方不仅都知道这一界线，而且都清楚没有其他地方可以作为双方都默认的容忍界线。如果前苏联人允许外人适度地跨越其边界

线，并且没有采取相应的大规模报复行动，那么，这将意味着，前苏联领土将成为导致战争爆发的诱饵。因此，政治边界通常成为停止前进的有效界线，但这并非表明边界具有法律强制性。边界成为双方默许的显著选择，是因为双方都期望寻求某种界线，而边界的独特惟一性使其担当了这一角色。边界是双方可以在某个区域划分的，并以显著的地理界线被双方接受的少数界线之一——或许是惟一一条界线，可以肯定的是边界线至少是这些备选界线之一，因为双方都可以清楚地看到这条界线。边界线具有明确的强制暗示作用，能够引起双方的共同关注，并否认了其他任何界线性—— 在缺少其他可辨认的选择情况下。

但是，政治边界或者国籍因素依然具有合法性，而且具有现实意义，我们有必要考虑有限化过程中其他重要的划分标准。在印度支那战争中，我们提供了大量的武器装备，但没有提供人力支援；在希腊的游击战争中，我们向希腊政府提供了武器装备、指挥人员和咨询意见，但是我们依然没有派驻美国军队。有人认为我们应该在印度支那战争中向法国人和南越政府提供空中支援，这样就避免了派驻地面部队导致中国人和前苏联人卷入的危险。

一位经济学家认为——与前面那位认为高精确的小型核武器是一种火炮的学者具有同样的说服力——军事装备和人力在军事行动中属于可替代资源，空中干预与地面干预在实际中没有本质差别，军事人才对于缺乏指挥人员和作战策略的军队而言尤如大腿肌肉一样重要。但是，有关现代武器的使用功能的再定义以及军事保障的机动方式有效性的争议表明，空地部队或海地部队的区别仅仅在于传统不同，除此之外没有其他区别。尽管如此，其中最关键的问题就是传统因素在有限战争中所发挥的作用。

实际上，我们对有限战争的分析研究涉及的主要问题就是传统。在研究过程中，我们主要分析先例、惯例以及暗示效果等因素。此外，我们的研究还将涉及到习惯法理论——包括避免相互破坏行为需要集体制裁以及个别制裁将导致规则的破坏，而规则的破坏将可能导致双方间有限条件失去原有价值甚至荡然无存。更严重的是，这一切的后果将使国际“权威”受到严重挑战，其他有效规则的权威性受到极大削弱。

核武器之所以与众不同，是因为强大的传统观念认为核武器具有独特的作用。我们没有禁止弓箭（尽管它们与核武器一样都能给人类带来伤害）的原因在于，人们传统上已经习惯于使用弓箭，并共同认为弓箭方便适用。然而，人们并没有使用核武器的传统。同时，人们传统上不认同核武器的使用——尽管已经具备了使用核武器的能力，而且使用核武器对自己有利，人们也一致认为现实中不应该使用它。

传统和惯例并不是有限战争的简单模拟或其中奇特方面；相反，传统和先例或惯例是有限性的本质。有限战争中的任何有限条件的根本特征就是双方共同承认，并被赋予权威性的心理、智力或社会特征。这些特征的权威性来源于双方相互认可的完全认知或“默式谈判”。某个有限条件之所以赢得一定的权威，主要原因是双方都担心否定这个有限条件后，无法找到其他合适的替代条件。有限条件暗含的基本原理是合法的并充满诡辩色彩，但缺乏法律、道德和心理支持。这些有限条件或许反映了合法的和心理的差别或者道德的差别。总之，这些有限条件说明，一定存在某些因素赋予它们自身的独特性和本质特征，并且为双方实现预期共识提供了依据。这个权威性存在于预期判断本身，而不存在于预期判断涉及的客体对象。

随着核武器的不断发展和改进，定义有限使用而非禁止使用核武器的有限条件的可能性日益受到人们的质疑。但是不管怎样，人们逐渐认识到，核武器的使用和效应存在等级性。但是，人们还是认为核武器不存在使用方式的差别，如运载方式或打击目标等，结果，在外人看来似乎自然不存在核武器有限使用的可能性。假设我们质问，如果愿意的话，我们该如何确定核武器的使用规模、运载方式、使用条件以及打击目标的有效性？答案是——从纯粹的技术角度——我们可以随心所欲地划定界限。事实上并不存在强制规定某个特定点一定为分界点，其他点则不能，或许这也正是很难发现分界线背后规律的重要原因；也不存在所谓的核武器使用程度、规模和范围，这如同其他为双方提供预期关键点的程度、规模和范围因素一样似是而非。合法的有限条件必须定性和非持续性的，而非定量和持续性的。问题的焦点不是为了更容易判断违规行为或强制双方的指挥官遵守协约，而是需要一个稳定的有限界线作

为暗示符号，一旦双方突破了这一有限界线，就很难再找到其他合适的替代选择。

有限条件通常出现于默式行为或谈判过程中这一现实，极大地增加了具有独特性和惟一性的显著条件的必要性。这些条件通常在博弈双方默不作声的斗争中实现，而不是出现在直率的沟通谈判中。但是，如果双方必须在缺乏直接沟通的情况下进行“默式”谈判，那么，某个特定界限必须能够在本质上区别于其他有限条件的连续统一体；否则双方都无法确定对方能否与自己承认同一个界限。当没有其他自然、可行的“显著”点或线帮助二者实现预期判断的共识时，只能以具有本质独特性和惟一性的某个纬线、国际时刻线或北极充当这一角色。

我们现在以核武器为例，通过下面的小游戏对上述观点进行检验。这里，我们需要合作争取某个奖赏：首先，我们坐下来，在没有任何事先约定的情况下，分别在纸上任意写出自己期望的核武器使用限度。这些限度应该包括我们希望看到的有关参数——如武器的规模、使用方式、使用者、使用频率、脏弹还是干净核弹、用于防御还是进攻、战术使用还是战略运用、是否打击城市、是否预先通知对方——这样可以清楚地看到我们所写的有限条件是否存在一致。如果我们所列的各项有限条件非常吻合，那么我们都将获得奖励；反之，如果不同，大家都将空手而归。通过这个合作获奖的小实验，我们可以清楚地看到大家能否在默式条件下就限制条件达成共识以及——对那些努力协作达成共识的参与者而言——大家在什么限制条件上存在分歧。我们允许参与者随心所欲地做出选择，包括核武器的各种有限使用，甚至包括两个极端选择：无限制地使用核武器和彻底拒绝使用核武器。

通过上面的实验，我发现某些特定限制条件——简单、间断、定性的“显著”限制条件——有助于选手做出同样的选择。同时，我推断那些选择其他限制条件的选手很难找到或者根本找不到志同道合者。（鉴于我们进行上述实验的初衷是实现合作，我们没有考虑被提议的有限条件的其他作用。实验中，选手选择某个有限条件的主要依据是为了与其他人选择同样的条件，并知道其他参与者进行同样的努力，只有这样大家才能双赢）。

尽管如此，我无法断言上述实验能够证明什么样的有限条件具有稳定性和权威性。但是，实验表明有限条件的某些特征——特别是他们的简单、惟一性、间断性、定性定义的质疑等——具有客观意义，至少说明有限条件与默式谈判过程存在密切联系。同时，实验还表明某些有限条件能够被博弈双方所期望的，成为达成共识的必要条件，并被双方承认与其他选择在本质上存在差异。

我们可以从上述讨论中得出一系列结论。首先，核武器与常规武器存在与有限战争过程相关的显著区别。与此同时，我们能够通过一系列手段强化或弱化、阐明或模糊这一区别。假设我们的言行严格符合传统要求，那么，我们就能借此强化传统的作用，增强上述区别的符号意义；反之，我们通过强调"核武器仅是另类武器"的观点以及肯定核武器的使用价值可以弱化这一区别——但是我们并不准备这么做。我们采取哪一种外交政策完全取决于我们是否将核武器与常规武器的区别看做是与前苏联人共享资源以及一个能够减少暴行的有效区别或传统——而非麻烦事、宣传义务、外交障碍和阻碍我们果断行动和代表行使权力的绊脚石。尽管如此，那些主张使用核武器以赢得优势和满足军事需要的人必须承认核武器与常规武器存在质的区别。只有这样，我们才能采取行动暂时消除这一区别。

问题不在于亚洲的中立国或欧洲盟国是否认识到这一区别，而是我们美国与前苏联两国如何处理相互间的关系——不管我们承认与否，美苏之间的确存在某种程度的谅解。而且，问题的关键是前苏联人是否认为我们与他们就限制使用核武器存在默契。为了实现战争的有限性，我们期望前苏联人或者中国人不要把我们在区域战争中首先使用核武器，看做是对限制使用核武器整体思想的挑战以及我们不受任何有关限制条件的约束；我们应该使他们相信，我们首先使用核武器并非违背有限战争的原则，以及我们仍然愿意与他们就有限条件的发现和承认存在默契；此外，我们必须避免别人对我们首先使用核武器产生其他不必要的释义。因此，如果正如我所分析的那样，假设核武器与常规武器的确存在差异并与有限战争存在联系以及我们追求使用核武器的最大自由，那么，我们应该竭尽全力破坏或消除这一区别。（例如，在有些国家，特

别是发展中国家，大量掘土计划主张率先广泛地使用“核爆破”无疑有效地削弱了核武器和常规武器的区别；类似的情况还有帮助发展中国家为了实现同样目的训练部队，如何使用现有武器在核爆炸中生存。）相反，假设我们愿意加强与对手的默契互谅，促使双方共同认识到核武器仅仅是另类武器并制定相关的保留条款，那么，我们就必须达成有关的暂停核试验条约（至少是对相关条约进行广泛讨论）。[1]

其次，一旦有限战争一方率先使用核武器，则相关的禁止使用核武器的规定将荡然无存。我们很难想像，一旦核武器首先应用于战争后，双方就核武器达成的默契能够再次在另一场有限战争中发挥强大作用。因此，我们或许无法忽视核武器与常规武器之间的区别，并将核武器应用到特定战争中，即使这样做对我们有利。如果是这种情况，我们必须将希望寄于二者之间的区别，以实现双方都放弃使用核武器的念头。如果我们抛弃传统，创造一个反面先例，那么，一场潜在的有限战争可能导致此前存在的相关默契消失殆尽，从此再也无法建立互信互谅的传统。（或许应该重新检验此前重视的某些限制条件或至圣思想，这样我们才知道，这些条件或思想是否是假设的禁止核武器条约的副产品，并将与禁止核武器条约一起消失。例如，我们或许应该重温前面提到的海军舰艇的作用，一是为了试探对方的反应，另外也是为了避免误解对方的真实意图，因为如果核武器被应用到战争中，对方可能采取不同的态度。）

我们得到的第三个结论是，一旦对方在有限战争中首先使用核武器，我们应该或者至少应该关注自己此前建构的行为模式和先例，以及我们认为的“核作用”和有限战争的初级目标。例如，假设台湾国民党政府使用核武器防止中国大陆的进攻，我们则应该更加关注核交流的特点而非其应用后果，即这一历史先河的建立以及我们和中国大陆在这一过程中发挥的作用。我们不应该仅仅将这一案例视为核武器应用于有限战争的特例，而是改变了未来有限核战争的模式。（假设一个学生手持

[1] 有关临时协议象征意义的讨论请参阅基辛格的文章，“Nuclear Testing and the Problem of Peace,” Foreign Affair，37:1~18（Oct.1958） especially pp.12~13。

一把弹簧刀威胁老师，那么，不管最初争议的焦点是什么，老师的第一反应一定是应该如何处理面对的威胁。）

最后，我们认识到——至少在核武器在有限战争中出现的第一时间——双方都同时身处至少两场不同的有限战争中：一场有限战争是为了实现最初的目标；而另一场则是双方有关核武器作用的默式谈判或博弈。为了进一步说明这一点，我们可以将核武器的使用与金门问题联系起来。通常人们总是认为我们会为了保卫金门而不惜使用核武器。但是，鉴于中国人或者前苏联人也可能使用核武器进行报复，我们或许无需考虑他们对使用核武器的看法会对进攻金门产生什么影响。因为对他们而言，更重要的似乎是如何回应我们首先使用核武器的行为本质。如果他们的核力量无法占有优势，那么，他们努力的目标将是如何与我们保持核均势而非核劣势。我们的摊牌既可能是为自己赢得完全胜利，也可能是彻底失败，因此，除非我们做好了充分的心理准备，否则我们必须就相关有限目标——核优势、核应用的传统和先例、未来战争的共同“规则”等——以及有限战争的其他目标与对方进行“谈判”（通过实际行动）。

附录 B

博弈论的非对称性

附录第一部分的结论表明，可能根本不存在纳什、卢斯和雷夫等分析的“无行为”（moveless）谈判博弈；[1] 或者即便存在，类似的谈判博弈也将具有前人始料未及的独特特征。这一论点的出发点在于“协作”的可行性含义——一个尚未确立的概念。本书第二部分的结论认为，“理性预期”这一概念并不能支持谈判博弈解决方案涉及的对称性因素的合理性。正如前面所言，这一论点的出发点是非理性预期在具体操作过程中的统一。

收益矩阵并不能定义显式（“合作”）非零和博弈——谈判博弈，因为博弈选择的操作实施必须具体明确。在一般情况下，达成协议的过程中的自由沟通或者“约束性协作”思想都反映了这些博弈选择的具体实施。因此，通常有关选手双方一旦达成相关协议即可分配 100 美元或者

[1] 参见 Nash，“The Bargaining Problem,” *Econometrica*，18:155~162（April 1950），“Two-Person cooperative Games,” *Econometrica*，21:128~140（January 1953）；约翰·哈桑尼，“Approaches to the Bargaining Problem Before and After the Theory of Games:A Critical Discussion of Zeuthen's Hichs' and Nash's Theories,” *Econometrica*，24:144~157（April 1956）；R.邓肯·卢斯和霍华德·雷夫，《博弈论与决策》（*Games and Decisions*，New York，1957，pp.114ff）。

双方可以就相关事宜进行完全沟通的观点都充分体现了博弈的含义。[1]

这类博弈场景通常都具有对称性的行为结构，尽管某些情况下的收益分配具有非对称性。选手双方平等地拥有选择沟通、拒绝对方的条件以及达成协议的权利。假设选手双方没有就 100美元达成分配共识，而是就收益矩阵范围内的某个值 X 或 Y 达成协作，那么，收益函数将具有非对称性，尽管行为结构是对称的。为了强调这一点，哈桑尼甚至特别明确假定了对称行为的存在："谈判双方必须遵守同一（对称的）行为规则。（不管他们遵守的是同一的理性行为法则或是屈服于同一的心理定律。）"[2]

我所关心的是在关注谈判过程中的"合法细节"的同时，分析有关博弈行为结构完全对称性假定的"协作"概念。此外，我们还需要关注什么是"非协作"。鉴于一个定义明确的博弈必须具有自己的终止规则，因此让我们首先分析博弈的终止规则。[3]

假设我们意在通过贴现的方式人为地避免我们的收益矩阵出现新元素，那么，我们必须期望博弈场景在利率因素出现之前终止。除了合作问题本身以外，我们并不想过多地考虑达成合作协议的时间问题。其中原因并不仅仅是为了易于操作，更重要的是除了在某些特殊情况下，博弈场景通常并非缺乏"无行为"，除非我们约定博弈为"无行为"活动。一旦选手的时间选择并不符合持续统一的贴现率标准，博弈将随着时间的变化而改变，选手也可能在实际中以无法达成合作协议而改变了博弈。尽管持续统一的贴现率被看做是一个必要条件或许多少有些例外，但不管怎么说，这并没有在我们实验的博弈模型中产生明确的假定因

[1] 实际上，卢斯和雷夫参照收益矩阵和下面的三个假设定义二人博弈：（1）博弈一方的预演信息必须准确无误地传递给对方；（2）博弈选手达成的所有协议均具有强制约束力，并受到博弈规则的保护；（3）博弈一方对博弈结果的评价不受预演谈判的影响。《博弈论和决策》，第 144页。

[2] 约翰·哈桑尼，"博弈论产生前后的谈判方法：有关朱森、希茨和纳什理论的重要研究" *Econometrica*，24:149（April 1956）。

[3] 虽然这里讨论的博弈场景是人为设计并非现实存在的抽象模型，但是这一模型有助于验证在虚构的抽象模型中，假定行为结构完全对称以及将非对称性看做是特例、对称性普遍存在是否具有积极意义。

素。因此，我们完全可以认为博弈已经一劳永逸地终止了。

或许终止博弈的最简单方式就是事先约定在某个特定时间敲响警钟。当然，除此之外，还有很多其他方法，例如让裁判员每数分钟摇一下骰子，不管什么时候，只要裁判员摇到某个矩阵函数，就宣布博弈比赛结束。(我们也可以假定,双方之间经过数次讨价还价后仍未达成协议时,博弈比赛结束。但是这一假定存在一定的风险，因为假定选手讨价还价意味着选手间“现实沟通行为”的存在，从而改变了博弈比赛的特点，使博弈场景发生了不同的变化。或许这将把我们引入到有关条件穷举的策略研究。)

简言之，提前假定博弈比赛在某个特定时间终止。为了简易起见，我们称其为“午夜”，也是为了易于操作。如果选手能够在午夜钟声敲响时达成协议，那么，他们即可按照协议的规定分配收益；否则，选手双方都将两手空空。

下面，我们该如何解释“协作”问题呢？为了方便读者理解，我们首先假设选手双方都拥有各自暂时的“正式”条件，并被裁判记录。当警钟敲响时，这些条件必须呈现给裁判；或许双方可以将自己的条件写在一个对方也能够看到的黑板上；或者双方将各自的条件装入一个密封的信封内，当钟声响起时交给裁判；或者双方将各自的条件敲入裁判室内的键盘。当钟声响起时，黑板上的结果被拍成照片，或者信封被递交给裁判，或者键盘被锁定，这样裁判只需要检查双方在午夜时提出的“即时”条件以确定二者条件是否一致，以及存在合作的可能性。如果双方的条件一致，那么，他们即可以以达成的协议分配现有收益；如果双方提出的总要求大于现有的分配收益，这将意味着双方没有达成协议，那么，二者的收益均为0。（现在假设双方提出的总要求小于现存分配收益，我们暂不对双方的选择结果进行要求，不管他们是按照达成的协议得到自己应有的收益，还是因缺乏共识而一无所获，则选手双方能否在午夜之前达成完全协议——即双方在午夜之前实现选择一致——结束博弈将变得不再重要。）

尽管从协议实现和存档方式的角度，我们能够以多种方式释义“协作”。但是，如果仅仅从完全对称的行为结构角度，我认为有一个特征应该引起注意，即博弈选手双方一定拥有某些时间用于选择或更改以前

的选择。（为了易于操作，这里我们假设做出选择和更改选择必须按照同一程序，这样我们可以假设一直存在一个“既有选择条件”。）对博弈双方来说，在午夜钟声敲响之前的有限时间内，一定存在某个关键时刻作为最后的选择机会。在午夜钟声敲响之前的这一时刻消失后，博弈选手再也没有机会更改自己的选择，重新提出新的条件。根据博弈的规则和保持理性的基本要求，选手双方都明白这一问题。同样，这一时间的对称性同时适用于选手双方。

此外，还有一个重要特征需要引起我们的注意。在一方选手不知道对方的最后选择时，他做出的最后选择往往是自然、合法的选择；同时，一方做出的最后选择应该完美无瑕，不要在博弈中给对方留下任何反驳的机会。在这个重要时刻之前，任何选择都可以更改；而在最后时刻，双方选手都有更改既有选择条件的权利。在完全不知道彼此选择的情况下，双方做出的选择都将是终极选择。[1]

但是，假设一方故意泄露自己的最后选择给对方，从而使对方有时间进行相应的选择；或者一方知道了对方的最后选择，并进行相应的选择，那么，双方的最后选择都不是终极选择——双方也知道彼此的选择不是最后选择。[2]现实情况也是如此。

尽管如此，现在我们还是得到了一个有关行为完全对称的谈判博弈的重要结论。这一结论是，在某些关键时刻，完全合作博弈有必要让步于默式（非合作）谈判博弈。博弈选手对此也是心知肚明。

最能体现博弈场景特点的不是选手双方必须在钟声最后敲响时达成公开协议或共同放弃收益，而是在某个关键时刻（或者是“特定时刻”）——钟声最后敲响之前的时刻——达成公开协议，或者在同一博弈中以默式方式达成协议。

[1] 巧合的是，只要我们坚持对称行为规则，即当结束的钟声敲响时，双方都可以“平等地改变自己的选择”，那么选手可以随时改变自己的选择完全不影响这一结论的效力。

[2] 这只是一个机制假设，即在重新选择的过程中，一方可以坚持原有的决定，也可重新选择。但是，如果选手是在午夜前一分钟或半分钟做出的选择是最后的选择，那么，这一情况将变得更加复杂，因为这一选择过程不可能重新开始，除非最后一分钟过后，关键点也随之过去。在后面的论述中，我们还会涉及这一问题。

选手双方应该知道这一情况，而且如果愿意的话，他们可以主动避免以公开方式达成协议，而选择以默式手段。因为，如果我们假设（暂时）默式博弈具有准确明晰的有效答案，那么，选手双方在博弈初期即可选择完全的最小最大行为战略。双方可以通过人为的回避公开协议来达到默式结果，直到警钟响起。假使博弈选手都非常理智，那么谁也无法通过唇枪舌剑从对方得到好处。

根据上面的观点，合作博弈的解决方案必须与默式博弈的解决方案相同（假使后者的解决方案明确有效）。事实证明了这一结论的正确性，因为默式博弈是合作博弈的必然结果。

从这里我们看到，博弈的合作特点消失得无影无踪。选手不一定非要等到 11:59 才出现，实际上他们根本不需要出现。前面为实现强制协议进行的沟通和努力似乎也已显得多此一举。事实上，根本就不存在与默式博弈泾渭分明的合作博弈。 [1]

当然，我们也不能保证这一结论一定准确无误。首先，默式博弈或许没有一个明确有效的解决方案。 [2] 而且，从显式博弈可能影响默式博

[1] 纳什在其 1959年的文章“二人合作博弈”中引入了一个博弈场景。该博弈场景在其后期呈现出明显的默式特征。这一场景与合作博弈的关系给了我们很大启发，有助于我们发掘在相关的合作博弈场景中影响选手“理性预期判断”（暗示的结果）的各种因素。作者认为，如果对称性行为结构得到严格的遵循，上述的关系就应该是机制问题，而非理智问题。同时，作者还认为严格意义上的对称性很难或者根本不可能准确地定义我们研究的终极课题——相应的非默式博弈。

[2] 需要强调的是，取决于清晰可见的零点的谈判博弈解决方法（纳什和哈桑尼提出的解决方法）——即选手无法取得共识时得到的惟一结果——无法适用于以选择矩阵为基础的合作博弈场景。根据规则规定，这样的矩阵没有零点（除了对角线外，所有收益均为零）结果，不存在一个“标准形式”既包含凸区域，又与零点相联系，除非存在一个完全充分的理论“解决”默式博弈问题（以选手重视的方式操作）。或许有人根据卢斯和雷夫的观点将徒手的“安全底线”（最大最小值）作为零点。但是由于存在仲裁因素和以假设为前提，所以，选手双方的默式博弈只能以这种方式取得成功。假设前提，特别是包含有效的完全战略点的假设（如布雷思韦特理论和后面脚注中涉及的卢斯和雷夫理论），很容易被实证所推翻，因为其前提是理性选手在缺乏沟通的情况下无法协调战略选择，尽管他们在面临选择冲突时可以有所举措。

弈这一角度，合作博弈具有的某些具体条件也能够发挥积极作用。同样，对选手双方没有任何强制约束力的前期沟通或许也能影响默式博弈的特征，例如下面的合作博弈的一些变化。

与其说选手双方根据双方达成的彻底分配协议分享了一系列成果，不如说他们某种程度上根据达成的协议分享了一系列成果，即他们仅仅可能根据双方在警钟敲响时达成的协议分摊现有成果。例如，假设现有100件物品，要求选手双方必须在警钟敲响时达成分摊其中80件的协议，如果二者达成了分摊这80件物品的协议，那么，他们可以根据这个协议得到各自应得的部分，剩下的20件则只好被送回仓库。[1]

现在，假设在显式谈判（合作）场景中，我们得到的结论是合作博弈存在有效的解决方案——选手双方必须达成一个完全协议——那么，我们或许应该认为有关这个问题的重新阐述是不合逻辑的。这一新论述认为，实际中的谈判应该是谈判双方将各自的要求和条件写在纸上；双方需要妥协和让步的部分应该包含在此前所列的条件中，必要时可以将这些条件删除。当双方列表中的各项条件没有冲突时，双方也就达成了完全共识。但是，当我们回头再看看默式博弈时，我们会发现这一阐述完全改变了默式博弈。改变后的默式博弈具有一个不合法的动机结构。双方都希望独自占有全部的成果，而不是和对方进行必要的分摊，彼此都知道对方怀有这一动机。双方谁都不愿放弃自己的要求，因为为争取剩下成果而付出的成本远远小于各自为息事宁人所付出的代价。这样一来，惟一一个平衡点也没能给双方带来什么收益。因此，与以前相比，更改后看似不太符合逻辑的新博弈场景发生了巨大变化。尽管如此，如果我们没有将博弈后期出现的默式场景看做是主导影响因素，那么，这

[1] 当选手双方面对一个分配对象时，如货币，那么相应规则将是他们必须根据双方在房子消除“交叠”（overlap）后做出的选择分配。一方的所得和对方的默许让步相同，例如一方在博弈最后阶段要求得到全部的65%，而另一方为55%，那么，第二方让出了35%，第一方让出了45%，这些数额即为争议之外的部分并构成“协议”部分（下面的脚注中有所体现），零点的潜在不明确是哈维·瓦格纳（Harvey Wagner）和约翰·哈桑尼所讨论的问题，“Rejoinder on the Bargaining Problem” *Southern Economic Journal*，24:280~482（April 1958）。

一变化也不会如此明显。[1]

现在让我们看看另一个例子。假设现有 100 件物品需要分摊，尽管就价值而言，这些物品可以在双方之间交换，他们还是必须就每件物品的所有权达成共识。假设博弈规则要求双方必须达成完全的分配协议，那么，双方在默式博弈中不仅要以合作方式分配物品的价值，而且还要以同样方式按件分配全部物品。如果情况这样，假设一方要求的一件物品的价值占全部物品价值的 80%，而对方拒绝其要求，那么，前者在默式博弈中就占有主动权，因为惟一能够分配这 100 件物品的方案就是满足前者的要求，使其占有全部物品价值的 80%。除此之外，双方协作达成其他分配方案——无论公平与否——的可能性将非常渺茫，尽管不公平，双方也将迫于必须达成协议的压力而只能接受前面的分配方案。由此可见，前期沟通仅仅具有战术意义；一旦博弈进入了默式阶段，沟通只能影响双方的协作方式。

根据上面最后一个观点，假设我们现在坚持对称行为规则，那么，我们得到的结论只能是，一方可能滔滔不绝地向对方解释自己立场的合理性，不给对方留下任何反驳的机会，而对方则只能目瞪口呆，因为双方都知道如果二者说出彼此想说的话，那么，就没有人愿意洗耳恭听了。换言之，完全对称行为这一先验结论似乎阻止了任何行为在沟通阶段丰富博弈的内涵。

但是，我们现在已经彻底将完全行为对称博弈远远地抛到了一边。[2]

[1] 这里，我们似乎能从分析中得出其他有利结论，即为了建构“真正的”合作（非默式）博弈和确立协议的规范定义而进行的分析研究将必然扭曲最终的默式博弈，因此，选手双方必须在警钟响起或遭受全面损失之前达成强制性协议。现在，选手为了在钟声响起之前达成协议，必须定义自己的“协议”概念。假设他们的定义与我们先前的定义类似，那么，他们需要做的是尽快将破坏后的合作博弈还原到完好状态，并且需要将新的博弈场景缩短一分多钟，而如果是默式博弈则需要缩短两分多钟。

[2] 有一个与前面脚注相关的细节或许值得深入探讨。假设一方需要花费一分钟提出条件或修改提出的条件；（与以前的版本相比）记录一个新条件的过程一旦开始就无法停止，除非整个过程结束。按照这一程序，双方在博弈结束前一分钟提出的任何条件都将是终极条件。如果这一终极条件在博弈结束前没有成功地传递给对方，则意味着博弈本质上没有变

考虑终结博弈和定义“协议”的方式等因素，我们在以后将对这一博弈场景进行具体详细分析。尽管如此，我们在这里不禁要问，“不作为”（moveless）或“行为对称”（move-symmetrical）博弈是否值得研究。看似公正的行为对称博弈是与“特殊场景”不同的普遍博弈，还是一个失去合作博弈最有趣部分的特殊的有限博弈场景呢？

在这里，我们必须强调，对博弈对称性的质疑并不意味着非对称性（asymmetry）假定一定成立，而是缺乏传统特征的对称性（nonsymmetry），因为后者没有肯定二者为既定结论，允许二者存在的可能性。

我们下面要分析的这个案例或许对大家有所启发。假设在一条路的尽头有 100 美元，两个选手谁先跑到路的尽头谁将得到这 100 美元。其实这个技巧性博弈很容易分析：排除意外事件和随机因素外，100 美元无疑归于跑得最快的那个选手。我们能够预测出理性行为（跑）和比赛的结局（100美元归于跑得最快的选手）。当然，结果也可能由于偶然

化。“同时性”现在意味着选手在各自的最后一分钟出于实际考虑，在做出自己选择的同时都无法看到对方提出的最后条件，无论这个条件是在最后一分钟的什么时候提出。但是，如果一方将自己的条件写在一个可见的板上，并掩藏一分钟以供记录在案，那么，对方将能够在最后数秒内看到这一条件，尽管他已经来不及修改自己的条件，除非推迟一分钟。（或者假设一旦各自条件记录在案，选手将无法看到彼此的选择。）在这种情况下，如果双方不是在最后一分钟同时提出各自的条件，那么，较晚做出选择的一方无疑完全知道对方的选择，结果较易做出选择的一方只能接受前者的条件，因为这是他获胜的惟一机会，除此之外，别无选择。因此，我们可以说，如果第一个做出选择的选手知道对方在等待自己的结果，那么“第二选择权”显然已经变得毫无意义。我们现在以下面的方式描述一个博弈场景：博弈选手双方已经进行了 23 小时 58 分钟，现在只剩下一分钟；而且博弈规则允许选手在最后一分钟的任何时候做出各自的最后选择。在这种情况下，博弈场景提供三个战略供选手选择：(1) 假设一方等待另一方做出选择，并要求得到全部收益的 99%；(2) 假设双方同时做出自己的最后选择，并接受默式博弈产生的结果；(3) 双方都在等待对方做出最后的选择。如果是第三种情况，那么，博弈将继续进行。假设双方只能等待有限时间，则我们得到的战略选择是：等待一次，然后要求总数的 99%；等待一次，要求默式结果；等待两次，要求 99%；等待两次，要求默式结果等。那么，这里的博弈（包含选手在最后一分钟选择的所有战略的“默式超级博弈”）将是我们要求的博弈。如果我们愿意接受的话，那么，这个博弈的“严格意义的解决方式”应该包括似有可能的战略选择（无限期的等待）并以默式博弈的解答要求而结束。

因素提前出现，但是通常只能到比赛结束后才能揭晓，并且比赛开始时双方往往很难看到最终的结果。我们必须另外制定有关结果的规则，但是这并不影响游戏的结果和我们的分析。

假设许多人同时参加一个类似的比赛，参赛者都知道每个人都会竭尽全力跑在最前面，那将出现什么场景呢？尽管每场比赛都有一个结果，但是其中发挥关键作用是只有有关结果的附加规则。既然每场比赛的结果已成定局，那么，他们为什么还自寻烦恼参加比赛呢？

一个典型的行为对称合作博弈就如同一场徒步比赛。在一定情况下，谈判如同徒步比赛一样徒劳白费。选手事先都知道所有的选择和战术都将没有优势，因为对方也可能拥有同样的选择和战术。如果行为完全对称以及双方选手对此的承认由于定义本身而被强加于博弈中，那么，我们能够注入谈判博弈的有趣因素也将变得平淡无味。

既然已经放弃了对称性这一假定，那么，我们怎么才能完善与发展博弈理论呢？虽然在博弈场景中通常存在许多行为选择机会，但是在实际博弈场景中，机会对选手双方而言总是不均等的。“行为”包括承诺、威胁、允诺，破坏沟通体系，寻求对允诺、承诺和威胁的惩罚，自我定位和真实情况的传递，意外因素对预期判断产生的消极影响，特别是当沟通匮乏的时候。有关这些“行为”的详细论述请参阅本书第二部分的第 5 章。

为了进一步说明这个问题，我们假设前面提到的合作博弈中有一个十字转门，允许选手离开，但是一经离开不得返回。一旦他走出这个十字转门，他的选择就将记录在案，直到钟声响起。现在我们有一个方式可以帮助双方做出“最终”选择或“承诺”：记录一个对自己有利并被对方知道的选择、先走出十字转门的一方将赢得最终的胜利。当然，双方都有可能赢得最终胜利。尽管如此，我们最终还是得到一个类似于跑步比赛的场景，离十字转门最近的那位将赢得胜利。通过分析这一战术和机制或实际安排，我们能够决定谁将赢得优势。

值得注意的是，我们在上面的博弈场景中引入十字转门并没有将战略博弈转化为技巧性博弈。实际上，一方只有在取得对方全面合作并阻止对方做出战略选择的情况下，才可能率先走出十字转门。如果仅仅是

走出十字转门，一方并没有赢得合法的实际胜利，那只是战略意义上的胜利。他迫使对方做出对自己有利的选择，这只是战略博弈的一个战术，尽管这一战术的运用依赖于技巧或地理优势。

现在我们可以在保持博弈完整性的同时，为其引入一定程度的对称性。当然比赛开始时，我们也可掷币决定谁距离十字转门最近，或者让选手双方处于类似的位置和速度，但是随机因素将影响谁先走出十字转门。尽管游戏现在实现了公正性，但是比赛的结果仍具有非对称性，因为双方都希望率先走出转门，而且保证现有条件对自己有利。[1]

同时，我们还要考虑其他可能出现的情况，例如存在两个门，双方可能同时走出大门。这些意外情况将使“对称性”成为一种可能，而非定论；而且如果行为和信息结构实际上有可能导致某种结果，但是僵局和预测结果成为可能。然而，由于我们关注的是缺乏传统特征的对称性，因此我们不必为某些可能性结果所困扰。

但是，如果一方能够提出条件并破坏沟通渠道，那么，他就可以通过提供双方必须接受的条件而赢得后续的默式博弈，因为双方在最后的默式谈判阶段必须达成共识。为了进一步说明这个问题，我们应该考虑当双方都有能力破坏沟通渠道，而且双方都意识到同时破坏沟通条件防止信息传递的必要性时，博弈将产生什么结果。但是，这种情况通常只是特例，而非普遍现象。

总而言之，完全“不作为”或“行为对称性”合作博弈并不是普遍存在的现象，而只是一个可能演化为普遍的默式博弈的有限情况。只有当“行为”得到允许时，合作博弈才会丰富多彩，意味深长。一旦双方之间的完全对称性包含在博弈的定义中，那么，行为的大部分意义将消失殆尽。如果引人入胜的是博弈行为，而非没有任何行为的博弈场景，则只有潜在的缺乏传统特征的对称性能使这些行为更加妙趣横生。

对称性通常不仅表现在博弈的行为结构中，而且还表现在其潜在结果的特征中或者实现结果的理性行为模式中。和哈桑尼（Harsanyi）的

[1] 有人认为博弈的预期价值仍然应该在双方选手间对称分配，从而使分析人员从平均结果判断博弈的对称性。但是实际情况表明，他这么做只能说明其对博弈和博弈方式缺乏深入理解。

理论一样，纳什的二人合作博弈也假定对称性为博弈内在因素。对称性的假定有助于双方找到博弈的解决方案，并将其局限于数学范畴。除了这一解决方案外，还有其他方案作为解决问题的基础与之相媲美。但是有关对称性的假定的合理性不在于其产生的积极结果，而在于违背了对称性就违背了选手双方的理性原则。这也是我希望批判攻击的根源。

我认为，虽然对称性与选手的理性原则相符，但是这并不意味着理性的定义包括对称性的同时，非对称性与徒手的理性原则相违背。因此，我下面要提出的观点对对称性解决方案十分有利。我认为，对称性是影响潜在结果的因素之一。

约翰·哈桑尼详细阐述了对称性与非对称性之间的关系。他说："至少在一种特殊情况下，谈判博弈具有明确结果，即完全对称的二人谈判。只有在这种情况下，双方才愿意与对方平分净利，因为双方都不希望自己的让步大于对方的让步。"[1]在后来的一篇论文中，他把对称定理看做是"基本定理"，并认为："这个定理给人的直观印象是，一个理性的谈判选手绝不会期望理性的谈判对手做出的让步大于自己的让步。"[2]

现在，这个以直觉得到的系统阐述涉及两个假定。首先，谈判一方的让步不会大于谈判对手对自己做出的让步。其次，如果是谈判一方以对方的立场预期判断自己的让步程度，那么，前提条件是他对双方间对称性的感知程度。

[1] 哈桑尼进一步指出："例如，大家都希望拥有同样的生产成本、企业规模、市场条件、资本和人力等条件的市场双头垄断者能够达成协议，平均分配利润。"

[2] 在此，我们有必要为大家提供完整的引述（经过作者的允许）："朱森和纳什的谈判理论主张将理性谈判选手针对对方谈判战略的预期判断具体化，如果双方都知道对方的效用函数的话。其实，朱森和纳什理论的基本假设是对称定理的存在。该定理认为，以数据形式表示双方最佳战略选择的函数具有同样的数学形式，当然，除非与双方选手有关的各种变量相互交换。直观地讲，这一假设为我们接受了一个定理，即在同一条件下，一个理性博弈选手不会期望理性的对手单方面做出更大的妥协和让步。"（Harsanyi，"Bargaining in Ignorance of the Opponent's Utility Function" Cowles Foundation Discussion Paper No.46，December 11，1957）。

这个来源于直觉，或者说是认真研究得到的系统阐述，即理性选手双方通常以理性的对手为参照物做出自己的预期判断的观点完全可以以心理学术语表述，但是一旦进行科学描述，就出现了问题。作为理性人，选手双方必须认识到惟一“理性”的预期判断是双方对博弈结果进行完全相同的预期。

作为对心理现象的描述，认为谈判一方总是希望对方妥协让步或接受自己某个条件的说法并不准确。对方让步或妥协的部分往往是对方希望另一方做出的让步或接受的条件。反过来，这也是第二方希望第一方希望第二方希望第一方所希望的。依此类推。为了避免在描述过程中出现“无限性”，我们只能说双方都感到彼此对结果的共同预测。双方之所以能够产生同样的预期判断，主要因为环境迫使二者必须做出同样的选择。对双方而言，这一选择具有必然性。实际上，双方都接受同一个权威——博弈通过影响双方对结果的认知能力来传递结果——双方期望的是彼此能够感知同样的解决方案。[1]

按照这些术语，哈桑尼假定的第一部分（明确部分）可以这么论

[1] 根据这一观点，在完全沟通条件下的谈判博弈中实现“理性预期”的思维过程本质上等同于默式博弈中实现协作选择的思维过程。尽管如此，实际解决方式可以存在不同，因为博弈的场景和具体暗示性条件或许不同。但是，两个解决方式在本质上是一样的，因为二者都取决于选手双方能否默契达成共识。更多事实证明了这一观点的正确性，选手在完全沟通情况下的博弈中取得的明确共识与选手在博弈开始前分别做出的预期判断相同。（假设在理论条件下，选手已经做出了预期判断。）在某种程度上，这类似于默式博弈。假设鉴于选手知道彼此的预期结果，他们都意识到双方会接受提前暗示的结果，则双方都会毫不犹豫地坚持自己的预期选择。

此外，这一观点的成立还需要一个条件支持。在双方充分了解彼此的价值体系后和即将分配同样收益的情况下，或许存在无限同样的解决方式，使选手双方得到同样的收益，但这并不妨碍选手克服分歧取得共识。默式博弈对选手双方提出了更高的合作要求，即选手必须做出相同的选择，即使是就收益的平等分配而言。有关同一区域界限划分的谈判与同时派遣部队抢占有利位置的情景还略有不同。尽管对土地价值的追求是一致的，但是要求或许不同，并有可能引起麻烦。因此，合作问题并非千篇一律。谁也无法保证默式博弈的解决方法（或者是非完全沟通或消息失调条件下的博弈场景）与显式博弈的解决方法一致。

述：在任何谈判中（效用信息完全），一方能够认识到的某个结果也被对方认为是一个暗示性“解决方案”。该假设的第二部分（隐式部分）是双方认识到的某个结果取决于数学结构的对称性。我们可以称第一部分为“理性解决方案”假定；第二部分则为“对称性”假定。

但是，问题在于对称性是否来自于选手双方的理性——预期判断的理智性——或者必须来自其他根据。如果是后一种情况，那么，这些根据是什么呢？其可信度如何？

首先，我们先来回答第一个问题，对称性是否来自于选手预期判断的理智性。我们可以将二者的理性放在一起综合考虑，并验证双方共同预测到的缺乏传统对称性特点的结果是否与理性假定相矛盾。假设双方都相信自己与对方预期判断到了某个同样结果，并且具有数学意义上非对称性，我们能否认为他们的预期判断是不理智的，与理性假定相矛盾呢？尤其是在下面的情况下：假设双方可以根据达成的分配协议分配100美元，协议规定一方得到80美元，另一方得到20美元，而且我们和选手双方都知道分配结果在一定程度上与效用成正比，那么，我们能否认为选手双方的选择是不理智的？

我们必须时刻小心谨慎以免将对称性赋予理性的定义中，因为这可能破坏理论内在的实证关系，从而简单地使对称性割裂为一个独立的定理。我们必须保证得到的理性定义没有对称性和非对称性在谈判过程中与定义相矛盾的痕迹。为了更好地说明这个问题，我们必须假设选手双方根据达成的协议将100美元分为80美元和20美元两部分，只有这样，我们才能确定双方在无法实现对称点的情况下可能出现的任何思维错误、误导的预期判断或不正当的个人私欲。

例如，B向A妥协，使后者得到80美元，那么B的让步是否正确呢？B或许预期认为——或者告诉我们，假设我们能够检验其准确性（如果利用充分的效用信息，这是一个非常适中的假设）——A会主动提出“要求”得到80美元；或许B猜测A期望得到80美元；或者B知道A知道自己会妥协，所以就满足了A的条件，接受20美元的分配收益。B知道A知道这一点；反之A期望得到80美元，并知道B已经做好了心理准备，因为B知道A相信B已经准备接受A的条件。反之

亦然。换句话说，双方都知道——他们后来告诉我们——彼此都知道分配结果一定是 A 得到 80 美元，B 得到 20 美元。由此可见，双方对彼此的预期判断都十分准确，不仅与自己的预期相符，而且也与对方的预期相符。我们或许为双方得到这样的共识感到迷惑，但是毕竟功绩在受到肯定的同时也不可避免地受到贬抑。一方面“理性解决方案”得到了完美展现；另一方面，博弈似乎也实现了双方期望的结局。尽管如此，如果我们还是无法得到同样的结果，那说明下面四个假设中必定有一个是错误的：(1) 理性解决方案的假定；(2) A或 B 判断是理性的；(3) 我们自己判断是理性的；(4) 我们分析的博弈场景与 A 和 B 进行的博弈场景存在同一性（在所有关键方面）。但是，根据掌握的情况，我们不能认为第二个假设——A 和 B 的理性选择——是错误的。

请注意：假设 B 要求分得 50 美元或者 A 要求分得 50 美元，二者都声称自己的选择是理智的，并就双方是否预期到同样的结果存在争议，那么双方可能都出现了“错误”，我们也无法根据掌握的现有情况判断谁的选择是理智的或者二者都不明智。除非我们将对称性引入理性定义，我们才有可能确定至少一个选手的选择是不理智的，或者理性解决方案的假定站不住脚。我们现在只有一个必要条件可以说明二者的非理性，而没有充分条件或必要条件说明单个选手的理智性。

如果我们质问他们如何做出预期判断，那么，我们将无法挑剔二者的错误。每一个符合要求的条件都能够成立，因为双方希望对方接受的条件都是自己的理智无法接受的。他们或许会为自己的选择寻找各种理由，如他们可能会说，黑板上的某个暗示告诉他们 A 应该得到 80 美元，B 应该得到 20 美元；或者说布告栏上写着 A 得到 80 美元，B 得到 20 美元。他们相信是这些暗示促使双方不约而同地做出了同样的选择——也可能是惟一可能的“预期选择”——那样，我们就无法抓住他们的把柄，证明其选择的非理性。虽然实际上，他们的选择可能是非理性的，但是我们没有相关证据证实这一点。

可是，这却成为别人攻击本人观点的把柄，因为我没有对双方的理性进行实验验证，并得到我上面提到的结果 80:20。尽管如此，即使这一结果存在，也只能说明双方非理性存在的可能性，并不意味着非理性

因素一定存在。然而即便这样，还会有人对此持反对意见，认为不可能出现这样的结果。争论的焦点是双方能否实现协作，做出同样的选择。下面将对此问题进行深入探讨。

假设双方都希望得到同样的结果，并相信这个结果也是对方所期望的，则他们一定会调动全部智慧寻求某种共同点。如果 100 美元转化为 9 999美分，双方必须就每个美分进行选择，而非 100 美元，那么，双方可能同时选择同一个美分，因为这是双方共同的预期选择。既然如此，双方如何才能每次都做出同样的选择，即双方每一次都产生同样的预期判断，尽管存在 9 999 个选择机会？答案一定是他们运用了某种技巧或某个暗示符号或合作约定。双方一定有意识或无意识中经过了一个筛选的过程，最终才做出同样的选择。一定存在某个关键点，促使双方将某个选择与其他选择区分开——即使他们没有意识到，至少我们的分析是这样。

现在，双方理性选手能够通过某种方式，而非仅仅领先巧合或某种神力将二者的注意力聚集在同一结果上，并“理智”地相信对方也会以同一个理由关注这一结果。如果情况如此，那么，双方又是如何做到这一点的呢？

答案是肯定的。正如第 3 章所论述的，他们能够做到这一点。双方可能充分利用现在的条件：暗示、符号或实现共识的规则。但是，上述条件中的任何一个都具有数学意义上的对称性。[1]

假设在一个没有内容，只有数学结构的博弈场景中，而且这个数学结构没有偶然因素是双方共同关注欣赏的，那么，我们除了依靠一堆数字外，别无选择。所有的数字都分为对称性和不对称性两大类。假如只

[1] 有关理性博弈选手的基本假设似乎是，成功战胜竞合需要某些规则的支持，而且不管是否具有理性因素，最好的规则都将是理性规则。例如，这一假设支持纳什的博弈模型将“非修匀”（unsmoothed）的默式博弈看做是平滑趋向结果 0的修匀（smoothed）博弈的极限。尽管非修匀博弈观点并非具有逻辑上的必要性。但是却具有较强的暗示作用并能够引起寻求共同选择的选手的注意，特别是没有其他更好的规律帮助选手实现共同预期时。这一有限性过程为选手在某个非修匀博弈中选择一个明确平衡点提供了暗示条件。当然，这一假设前提同样也支持其他平衡点的选择方法。

有一个数字属于非对称性，那么，对称性原则足以有助于双方做出同样的选择，而且可以建立一个严格定义上的博弈场景，排除个人因素或其他环境因素，即没有任何明确条件促使二者做出同样的选择，除非博弈本身具有缺陷。[1]

换言之，数学意义上的对称性能够实现双方做出同样的选择，因为这一对称性——承认博弈场景具有的特征，如双方效用体系的全部情况——为双方提供了实现共同选择的方式。这一方式能否实现将取决于现有的选择条件。

除此之外，选手还有很多其他方式寻求相同的预期选择，其中许多优于对称性概念的方法论在第 3 章中大量涉及。由此可见，我们完全可以建构一些博弈场景，其中数学意义上的对称性能够提供共同预期目标或者博弈其他其他方面也能提供预期的焦点。[这些方面不存在于博弈的数学结构中，仅仅作为“话题内容”（topicalcontent），即取决于双方徒手和战略的“标记”，例如第 4 章中提到的卢斯和雷夫的故事。]

我无法证明数学意义上的对称性，以何种方式或在多大程度上主导双方的“理性预期判断”。但是，我认为允许影响选手预期判断的不同因素与对称性因素同时存在，极大地改变了对称性假定的地位。假设我们相信只能通过收益函数的数学属性实现理性选手预期判断的协同，那么，对称性毫无疑问将成为无可争辩的定论，特别是当我们能够为对称性找到一种合适的定义，满足某些引人入胜的定理的要求。但是，如果

[1] 根据这一观点，纳什理论（追求最大效用结果的解决方法）表明，即使是在数学王国里也存在众多惟一性和对称性，从而为做出选择提供统一规则。由此可见，对有效标准的需求促使可行性标准（公理）足以产生明确选择。布雷思韦特（Braithwaite）理论可以以同样方式描述。两个解决方式存在冲突说明数学家也无法就哈桑尼假说提供统一和充分的数学支持，即协调预期判断，选择同一结果。（R.B.布雷思韦特，“Theory of Games as a Tool for the Moral Philosopher”，[Cambridge，England，1955]；同样，卢斯和雷夫的《博弈论与决策》一书也都涉及布雷思韦特的解决方式，145ff。）布雷思韦特将问题看做是一人仲裁问题，而卢斯和雷夫则为了强调纳什理论的灵魂——智力合作——从仲裁而非战略的角度重释纳什理论（第121~154 页）。一个合法性的解决方法要求惟一结果具有某种理性因素；当没有其他选择空间时，完全诡辩或许有所帮助。

承认其他因素——收益函数数学结构的非必要成分——能够取代对称性的地位，发挥其作用，那么，就没有充分理由假设对称性发挥了 99%或者说 1%的作用。对称性的优势也不在其数学特点，而必须重新内省其优势所在。更多的争论将局限于某些人对作为博弈选手的博弈理论家或者某些实证研究的特定问题的个人争议。

因此，作为战略理论的标准博弈理论虽然依赖于选手双方的预期合作，但是现在它遇到了一个依靠经验合作的对手。后者，人们如何协调他们的预期判断取决于二者制定的策略和身处的具体环境。一个理性选手必须使自己在某个情况下面临某种经验问题；当双方在博弈中意识到存在与对方达成某种共识的可能性时，他们或许能够通过默式协作实现相同的预期判断。双方理性的对称性存在的前提是，一定存在某种思维过程是理性选手无法胜任的，即双方必须以其他方式达成共识而非数学意义上的对称性，以及双方对这一点的了解。理性选手的实际行为是否违背了该理论的某些要求，以及是否应该忽视其中的一些根本原则，仅仅是个实证问题。[1]

[1] 有趣的是，在寻求默式博弈的对称性解决方法的过程中，卢斯和雷夫忽视了两个最重要的、最有利的坐标。他们在分析下面的矫形后发现，该矩阵的左上角和右下角存在完全战略平衡点。卢斯和雷夫之所以忽视这两个因素，主要原因是他们认为：“无论 Ⅰ 以对称性条件赋予 i 或 ii 多大程度的理性因素，选手 2 将得到同样的理性条件，因此无论怎么样，我们看起来都将必输无疑。（Ⅰ 已经取代了 i 和 ii的暗示地位。）”在此之后，他们又分析了一对最大最小战略选择，但都不满意，因为二者都缺乏平衡点，而且另一个最小最大战略效果更不理想。尽管如此，最重要的问题是充满理性和想像力的选手是否如卢斯和雷夫坚持的那样无能，

Ⅰ	Ⅱ
2 ＼ 1	−1 ＼ −1
−1 ＼ −1	1 ＼ 2

选手能否在缺乏沟通的情况下做出同样的选择？这是一个实证问题：第 3 章的实验为这个问题提供了明确答案，或者至少说明在某些情况下，答案是肯定的。当然，他们似乎很难做出非对称性的战略选择。但是，更艰难的是承认他们必须这么做，至于以何种方式则是一个具体操作层面的问题。双方必须共同努力，相互默契配合才能找到做出同一选择的暗示条件。当然，对选手而言，上面矩阵的解决方法并非公平。尽管由于完全意外或偶然因素，应该指

在很大程度上依赖于实验研究的内省博弈也说明这一问题。 假设代表博弈潜在结果的所有点分布于两个矩阵坐标的右上部位， 那么，不管我们对对称性假定和纳什解决方案的某个对称性多么偏好， 我们首先将“纳什观点”看做是显式谈判博弈的理性结果。 [1]下面，我们将对博弈场

责一方的收益可能会大于另一方的收益，但是我们必须假设理性选手能够接受这一不公平的现实，如果暗示条件要求这么做的话。只有不公平的平衡点才能帮助双方达成共识；否则，否定不公正性就是否定双方为了寻求胜过任何对称性结果的结果，而寻求并接受隐藏的暗示条件。卢斯和雷夫在分析后总结说：“尽管这个看起来公平的博弈场景具有一定程度的对称性，但是要具体发挥这些对称性因素却异常困难。”尽管如此，这一看起来公平的博弈的真正答案却具有某种程度的非对称性，特别是在某一特定场景中，目标就是如何发挥这些非对称性。详见pp.298ff。

[1] 在博弈选手对彼此的效用体系（主观价值判断）十分了解的情况下，纳什提出自己的博弈解决方式，即博弈结果必须将选手双方的效用体系收益最大化。假设所有可能的结果都隐含于一个曲线图，而这个曲线图的坐标是标识选手的效用收益，那么，博弈解决方案只能位于收益矩阵右上部分的某个点。（该点之所有具有惟一独特性，主要基于以下原因；假设存在两个点，那么这两个电将被一条直线所连接，而这条直线则表示混合各种随机因素或原来两个结果后可能得到的两个结果；直线上与他们相连的各点能够增加选手的效用收益。换言之，这一区域由于几率混合存在的可能性而凸曲，并具有惟一一个最大效用收益点（纳什点）。

纳什提出的这一“解决方法”最显著的特点在于，该解决方法不受选手双方间效用尺度的兑换率的影响。换句话说，我们对双方各自效用的评价并不影响这一解决方法。同时，这一解决方式还符合其他条件，特别是与形成对称区域的选手效用尺度有关的某对固定评价（或兑换率）的条件，右上方的中间点即为最后的解，也是选手双方间最对称的一个点。（这也是惟一一个满足所有指定条件的解答。纳什表示，任何满足其条件的解都将产生促使选手双方效用收益最大化的结果。）这里为了研究的需要，我们将对称要求看做是博弈解决方案的一般属性，并将其他条件看做是为了实现解决方案的惟一性而对对称概念的改进。详情请参阅本书有关纳什等人的论述。其他相关文章还有罗伯特大主教即将发表的“The Nash Solution of Bilateral Monopoly and Duopoly”，该文对纳什理论进行了诠释和批评。有关纳什点应用到仲裁理论的资料有莱曼 E.艾伦（Layman E.Allen）的“Games Bargaining:A Proposed Application of the Theory of Games to Collective Bargaining,” *Yale Law Journal*, 65:660（April, 1956）。

我们有必要顺便强调，纳什理论也需要一种方式能比较选手效用尺度——与选手个人的效用评价无关，并能够独立存在。然而，鉴于纳什理论将效用兑换率作为根本原则，因此纳什理论依附于选手效用的内在不可通约性（incommensurability）。假设选手的效用尺度原则上具有可比性，尽管操作起来比较困难，纳什理论看起来也并不是排除困难比较有效的手段；假设选

景的各个变量进行分析。

首先，我们以默式方式完成同样的博弈场景。双方从各自的坐标轴上选择一个值。如果得到的分数位于边框上或者边框内，我们就得到了自己选择的坐标的（效用）值。我认为在自己希望的思想框架内——这使我们在显式谈判博弈中更倾向于接受纳什的观点——我们应该接受纳什的观点。不要问为什么，让我们继续进入另一种博弈场景。这也是一个默式博弈场景。不同点在于，除非我们选择的坐标点正好位于坐标轴上，否则我们将一无所获；或者除非我们取尽所有的值，否则我们什么也得不到。在这种情况下，谨小慎微也不能帮助我们，一方必须每一次都与对方做出同样的选择。因此，我认为在目前的思想框架内，我们应该接受纳什的观点。

最后，我们考虑另一种博弈场景。我们在分析完一个刚刚完成的博弈示意图后，意识到我们现在是一对同甘共苦、荣辱与共的博弈搭档。同时，我们现在的博弈场景是以我们选择的谈判博弈场景为原型，无法进行沟通，坐标点正好位于坐标轴上。如果我们做到了这一点，那么，我们将同时获得奖励——不管收益如何，我们将得到同样的奖励——否则，我们将一无所获。鉴于此，我认为在完全合作博弈中，我们或许也应该（必须）在现有的思想框架内接受纳什的观点。

或许有人会问为什么。道理很简单，我们需要理性化因素实现一个惟一性结果。只有在这样的环境中，谈判类比才能满足这样的条件，除非可能存在一个凸点（即便如此，也将发展为纳什点）；或者当坐标线是一条直线或曲线时（最终回到纳什点），存在一个中间点；或者是明显指向某个点的启发性方式。否则，除非博弈场景存在缺陷（如界线上有一个点），我们才可能回头重新研究对称性的“惟一”定义。与纳什

手效用原则具有可比性，那么，纳什理论以不可通约性原则得到博弈的解决方法并没有多大的意义。现在，虽然人们有关博弈论和经济理论的认识依据与选手间的效用比较并不协调，但这并不妨碍仲裁思想与效用比较的密切联系。经济理论发现效用概念的应用有助于促使效用理论与决策理论的协调，从而得到经济决策理论的副产品“福利经济学”。假设有人放弃这些相关性，那么，就会有人为了仲裁的派生原则而以心理学或生理学方法衡量选手的“效用”或者依法建约某个比较规则——尽管具有偶然性，但是还是符合仲裁的社会价值。

定义的对称性一样，我认为的对称性同样具有可行性——虽然没有像其他对称性那样简单（如一个图表的 45°线与另外一个类似图表的线相交形成的对称性），但是至少在其复杂水平上简单易懂。

如果纳什点在谈判博弈中对我们具有极大的帮助，那么，情况一定更是这样，因为我们相信纳什点对对方具有同样的吸引力；反之，他们也相信我们会接受这一观点。在完全合作博弈场景中，纳什点作为一个独特的点吸引双方的视线，双方都明显地关注着这一点。

或许有人会问，这又能说明或者暗含什么问题呢？我并非想证明纳什的观点。我们只是认为，纳什点对博弈理论家（作为内省的博弈选手）的吸引力或许与我刚才得到的结论相矛盾。或许是纳什点在完全合作博弈中的核心属性——在没有非数理缺陷帮助的情况下，对定义惟一的对称性的准确运用——对默式和缺乏合作的博弈边线变量产生了决定性影响。反过来，这一属性又成为对默式要求不高的博弈边线区域变量的可靠向导，并受到显式谈判博弈选手的青睐，因为他们希望双方的预期判断在其他任何地方也能产生共同。

换句话说，通过假设对预期合作的需求，我们似乎为纳什定理找到了存在的理论依据。纳什理论需要的正是解决方案的存在这一假设。很显然，只有默式合作才能提供实证根据，（偶尔）理性预期判断也暗含某个特定（或许有效的）结果。这也不得不使博弈选手认为，在仅仅具有数学属性的博弈场景中也存在同样的可能性。

纳什理论正好验证了这一假设的合理性——如果纳什理论能够比其他数学解决方案更能体现数学的美感。最后的焦点只能是，数学方法的普遍存在并不意味着博弈论的普遍性。

附录 C

重释“非合作”博弈的决策理念

利益完全共同博弈或合作博弈反映了博弈论某些决策理念的合理性，特别是严格意义上的“非合作”博弈决策方法。通过“研究这些方法理念背后的各种关系”，我们能够分析为什么理性选手对这些理念方法如此青睐。[1]

	I	II
i	1, 1	0, 0
ii	0, 0	3, 3

图附录 C-1

图附录 C-1 表示具有严格意义上解决方案的默式博弈。（在图附录 C-2 中，解决方案是双方的第二和第三战略选择。）卢斯和雷夫分析

[1] “非合作”是人们对博弈的传统称呼，它表示在博弈过程中，选手双方没有明显的沟通行为。很不幸的是，“非合作”在某种程度上表明，博弈选手间沟通的匮乏正是博弈合作难以实现的根本原因，这或许不是“非合作”概念的初衷。正如第 3 章和第 4 章所述，在许多默式非合作博弈场景中，合作——博弈选手都十分重视的互动行为——无疑是非常关键的因素，甚至可以说是核心因素。

了类似解决方案的定义："非合作博弈在符合下列条件的情况下具有严格意义上的解决方案。（1）选手双方的战略选择之间存在平衡；（2）所有允许的平衡选择必须具有互换性或等价性。[1]

	Ⅰ	Ⅱ	Ⅲ
i	1, 1	0, 0	0, 0
ii	0, 0	3, 3	3, 3
iii	0, 0	3, 3	3, 3

图附录 C-2

一对平衡的战略选择是双方的最佳选择（或者至少与其他选择一样有利），一方的选择应该与对方的选择相匹配；双方允许的一对选择应该不受其他选择的影响，即该选择得到的收益不逊色于其他矩阵格内的收益。对选手双方而言，平衡选择是等价的，因为每一方都得到同样的收益；而且，如果双方依据平衡点做出战略选择，那么，平衡选择也可互换。（换言之，只有当双方战略选择具有等价性，其产生的收益才具有互换性和等价性。）因此，图附录 C-2 中的战略选择（ii，Ⅱ）（iii，Ⅲ）（ii，Ⅲ）和（iii，Ⅱ）都具有等价性和互换性，也就是双方允许的平衡选择。

在完成这个定义之后，卢斯和雷夫立即进行如下表述："第二个条件避免了多种平衡选择之间相互混淆。"（楷体字为作者所加）

简言之，问题的关键是如何避免混淆或歧义现象，这也是第 3 章涉及的合作博弈的核心问题。图附录 C-3 并没有严格意义上的解决方案选手双方的第二选择和第三选择不具有互换性和等价性——即在 4 个组

[1] 《博弈论和决策》，p.107f。这一特定博弈解决方法看似纳什于 1951 年提出的观点，但本质上则不同。了解其他几种博弈解决方法的区别，请参阅第 5 章的卢斯和雷夫的案例以及纳什的"Annals of Mathematics"，54：286~295（1951）。

合中没有一个是等价的。双方的战略选择除了造成一定的混淆外，并没有影响双方的实际收益。在图附录 C–1 中，双方知道自己必需的选择；在图附录 C–2 中，双方知道自己应该做出的选择；而在图附录 C–3 中，双方不知所措，不知道应该如何选择。图附录 C–3 中，双方之间沟通的匮乏导致二者都一无所获。由于缺乏有效的沟通，选手双方只有 50% 的机会赢得 3 分值，即各自的预期价值均为 1.5。

	Ⅰ	Ⅱ	Ⅲ
ⅰ	1, 1	0, 0	0, 0
ⅱ	0, 0	3, 3	0, 0
ⅲ	0, 0	0, 0	3, 3

图附录 C–3

为什么在图附录 C–1 中，默认结果是（ⅱ，Ⅱ），而非（ⅰ，Ⅰ）？一个简单的答案是，前者优于后者。但是，这只是全部答案的一部分。我们能够在图附录 C–4 中找到另一部分答案。图附录 C–4 基本上与图附录 C–1 相同，不同之处在于双方选择偏好程度不同。

	Ⅰ	Ⅱ
ⅰ	9, 9	0, 0
ⅱ	0, 0	10, 10

图附录 C–4

在图附录 C–4 中，对双方而言，问题的关键是如何在 9 或 10 和 0 之间做出选择，而非 10 和 9 之间。简单地说，两对平衡选择很大程度上具有等价性，但是缺乏互换性。尽管双方某种程度上都关注自己能否

得到 9 或 10，但是他们更关注的是能否得到 0。双方的共同利益在于做出明确的选择，避免出现“混淆”。

他们需要某种暗示、规则或启发以做出共同的选择。在图附录 C-4 所示的抽象的博弈场景中，只有数字能发挥这样的作用。在选择较大值或较小值之间，前者显然具有更大的吸引力和可行性。我们不禁要问（ii，II）比（i，I）多的 1 美元对双方而言有多大价值。实际上，这 1 美元价值不在于其货币价值，而在于其起到了暗示符号的作用。对双方而言，是 9 与 10 之间的差别促使二者做出了同样的选择。在图附录 C-5 中，我们假设选手双方没有找到合作规则，那么，二者的预期收益分别为 5。

	I	II
i	10, 10	0, 0
ii	0, 0	10, 10

图附录 C-5

（事实上，图附录 C-5 表示的博弈场景以图表形式更容易说明问题。但是第 3 章的实证结果表明并没有这样做的必要性，因为一个具体的矩阵允许左右、上下、前中后存在不同。为了分析的需要，我们必须假设选手做出的战略选择必须以这种形式，并表明理性选手在智力上无法实现同样的选择。一个完全简单的没有什么暗示的博弈场景，将缺乏明确的标识或具有对称收益。巧合的是，拥有众多战略选择的默式博弈也没有“完全”形式。战略的无限性只能以一个生成公式展现给选手，因为任何生成公式都为选手提供某种选择战略的方法。）

假设我们选择（ii，II）作为强调的结果，黑体印刷体指向定结果，或者有个脚注写着一旦出现混淆现象，管理人员建议选择（ii，II），那么情况将保持不变。选手需要的是实现战略合作某种暗示。如果他们无法在收益的数学分布图中找到这些暗示，那么，他们只能另寻高就了。战略或许能以这种方式出现，标签或者暗示为选择或区别这些

战略提供依据，理性选手也会发现大有裨益。[1]

附录 C 旨在强调严格意义上的解决方案的重要属性——理性选手选择某个战略的原因——具有重要的暗示作用，即默式沟通的一种方式；当沟通的缺乏非常严重时，这有助于理性选手实现默契合作。当然，这并非类似解决方案的惟一属性，但却是博弈选手选择战略的规律的重要组成部分。

在本书涉及的博弈场景中，我们也可以以另一种方式说明这一点。我们能够以某些沟通成本实现沟通，并分析这些博弈场景以求证这些成本的价值所在以及产生“解决方案”的沟通渠道。博弈选手可以充分利用本书提到的各种“暗示条件”，尽管选择什么样的沟通方式以及如何看待这些沟通方式完全是一个经验问题。正如语言认知和语法的要求有助于消除由于信息沟通不畅而出现的歧义现象，认知或剧法要求、诡辩论或几何学要求也有助于消除默式合作场景的模棱两可因素。

我们可以进一步发展这一观点。在图附录 C–6 中，我们假设双方选手由于智力的局限性无法做出其中的战略选择，特别是不能以某个正方形的形式存在或以数字或字母充当暗示符号，或者——如果带有标识的话——选手各自拥有自己的标识符号。如果是这种情况，假设所有标识字母没有任何暗示合作的作用，那么，解决方案将是战略选择（iii，Ⅲ），即结果为 9。这也是平衡点要求的最小值，而且在其他值存在混淆的情况下，这个值明确惟一并发挥了暗示合作的作用。仅就收益结构而言（即没有引入“标识”、预制矩阵或博弈的完全定量结构以外的任何因素），与图附录 C–7 相比，这个解决方案仍不失为一个好结果，尽

[1] 博弈对选手理性和智谋上的要求很像猜谜。猜谜在一定程度上要求猜谜者必须找到某种启示或暗示，但是这些暗示或启示必须难易适中。对猜谜者而言，寻找这些暗示条件应该既不至于难于上青天，也不应该易如反掌。（即当这些暗示或启发条件出现时，理性选手应该意识到。）从本质上讲，博弈与猜谜都属于二人问题。猜谜的解决方式取决于另一方隐藏于谜语中难易适中的暗示或启发条件。原则上，这些暗示条件的隐蔽性应该难度适中。猜谜者不能仅仅通过经验得到答案，也不能经过逻辑推理后得不到谜底。“暗示理论”（hint theory）是博弈论的内在经验部分。

	Ⅰ	Ⅱ	Ⅲ	Ⅳ
ⅰ	10 10	0	0	0
ⅱ	0	10 10	0	0
ⅲ	0	0	9 9	0
ⅳ	0	0	0	10 10

图附录 C-6

	Ⅰ	Ⅱ	Ⅲ	Ⅳ
ⅰ	9 9	0	0	0
ⅱ	0	9 9	0	0
ⅲ	0	0	10 10	0
ⅳ	0	0	0	9 9

图附录 C-7

管后者符合卢斯和雷夫的定义，而前者与之矛盾。[1]

图附录 C-8 和图附录 C-9 均无严格意义上的解决方案，但却说明了同一个问题。在附录 C-9 中，选手似乎就是否选择（ⅱ，Ⅱ）存在争议。一方认为在不知道选择（ⅰ，Ⅰ）还是（ⅱ，Ⅱ）的情况下，自己应该考虑最保守的选择。如果行错选了上面的行，那么，他将一无所获；反之，如果错选了下面的行，他将得到 5。这里的错误选择意味着一方无法与对方达成共识得到最佳结果 10。选择下面行的选手或许会

[1] 意在继续探讨这一观点的读者可以得到大量相关博弈场景的实证。

为自己辩护，自己这么选择至少可以稳得 5，而不必冒险选择 10。与结果 10 相比，自己得到结果 5 具有同样的意义。尽管这或许就是他按照自己的“理性”做出的选择，但是下面的推理过程似乎更具有说服力。

	Ⅰ	Ⅱ
i	10 10	0 0
ii	0 0	10 10

图附录 C-8

	Ⅰ	Ⅱ
i	10 10	5 0
ii	0 5	10 10

图附录 C-9

“在比较（i，Ⅰ）和（ii，Ⅱ）之后，我发现与对方无法达成共识。或许在其他某个地方存在答案，我们必须试试看。其他惟一可能的地方就是（ii，Ⅰ）和（i，Ⅱ）。这是否暗示我们能够共同选择 10？答案是肯定的。它们似乎不约而同地指向（ii，Ⅱ），并提供了令人信服的理由使我们相信或者至少有个借口骗使我们相信选择（ii，Ⅱ）比选择（i，Ⅰ）好，因为我们至少需要一个自欺欺人的借口使自己相信，一对平衡选择比其他选择搭配更有利、更卓著、更显著或者更符合条件，假如我们没有一个令人信服的理由的话。其中原因可能是，我们无法找到其他更好的规则或建议或暗示继续遵循，我们或许能够以这个规则实现共识。”

在这个博弈场景中，选手不会进行第二次选择，因为第一个选择的收益虽是 5，但是已经足够比一无所获好很多。因此，他们更倾向于选择 5，而且将结果 5 和 0 的分布看做是协调双方行为的某种暗示。对双

方选手而言，他们有必要——双方都意识到对方也意识到这一有利价值——注意 5 的位置，尽管这只是实现共同预期的第一步。在图附录 C-9 中，双方更倾向于共同选择（ ii ， II ）。这如同白纸上的矩阵箭头指向右下角的选择，而且这个箭头仅仅依靠暗示条件和双方协调彼此预期，并非依赖于逻辑推理或某种权威。[1]

冲突利益

现在我们开始探讨冲突与合作并存的博弈场景。图附录 C-10 和图附录 C-11 表示的博弈场景均含有博弈允许的平衡点，但是没有严格意义上的解决方案，因为平衡选择既无等价性亦不可互换。

[1] 假设一方选手选择了 ii 或 II，那么，我们有必要寻求辨别二者背后不同动机的可行性方法，即使仅仅是为了验证这一概念的可行性。我们或许可以通过以下方式发现两个动机（“保险”动机和“合作暗示”动机）的不同之处。我们可以将选手置于一个类似于图附录 C-9 的博弈场景。二者不同之处是，后者中的收益 5变为 0~9 的变量，其他收益值不变。然后，我们询问选手愿意为了我们付出多大代价而完全好博弈游戏——即选手愿意为与一个选手竞技和得到真实的现金收益的机会付出多大代价。（换句话说，我们问他愿意为优先选择不同的变量付出什么代价。）如果他的答案对某个正数收益值来说没有任何意义；或者尽管如此，他仍对这个博弈游戏注入极大期望，并为之做了充分准备，诸如随机战略预期值等，那么，我们可以得出肯定的结论：右上和左下格的收益值对于选手来说，将是吸引力最大的暗示符号。例如，假设选手愿意每次付出 9.50 美元参与图附录 C-9 的博弈场景（这同时意味着列将有 90%的概率选择 II）;如果将 5 改为 1，那么，他将每次付出 8.65 美元；或者将 5 改为 9，那么，他将愿意每次付出 9.95 美元；或者他愿意以每次 5 美元参与图附录 C-8所示的博弈场景，那么，我们得到的肯定结论是：对选手而言，右上和左下格的函数或值具有巨大的合作暗示作用。如果，该选手的预期成本表明，选择 I 和 II 就右上或左下格的收益值而言没有区别；特别是当他选择以算术平均做出选择，那么，这就证明了其保守的动机（请注意，客观存在的“右上格”和“左下格”只是作者采用的简化形式。对于我们研究的随机（*unlabeled*）战略而言，这样的简化方式不会对博弈选手产生影响。类似的战略选择没有出现在收益矩阵中，或者仅仅以“ i ”和“ ii ”为模糊标识。即使这些战略选择存在标识，它们也应该通过随机过程产生，并且与对方展示的立场或标识选择的随机过程区分开。需要注意的是，行一定不能知道列的收益矩阵是否是像图附录 C-9 那样，或者将左上格和右下格的值换成较大的收益值）。

	I	Ⅱ
i	6 4	0 0
ⅱ	0 0	6 4

图附录 C-10

附录 C-11 表示的完全以抽象形式存在的合作问题显然“无解”，因为战略选择没有对应的暗示符号。对双方来说，只能依靠侥幸实现博弈场景能够允许的（有效）结果。[1]当然，附录 C-11 或许存在解决的可能性，因为双方宁愿接受各自的“次优”平衡点，也不愿丧失合作的机会。二者的共同利益在于一起寻找达成共识的暗示符号。如果情况如此，那么或许有人会问，他们为什么不接受其他方格中指向（ⅱ，Ⅱ）的暗示符号呢？[2]

	I	Ⅱ
i	4 6	3 2
ⅱ	2 3	6 4

图附录 C-11

对于其中一方选手来说，这也许并不是最佳选择。但是，如果天上掉馅饼，乞丐难道还有选择的权利吗？是否还存在什么其他暗示符号呢？即使这些暗示符号没有积极意义，也会具有反面作用吧？例如，它们的箭头指向（ⅱ，Ⅱ）而非（i，I）时，双方完全可以把这个标志看做是某种信号，而非仅仅是箭头。当然，这或许并不能发挥多大的作

[1] 本书前面的脚注提到了另一个类似的收益矩阵，不同之处在于该矩阵没有完全抽象的前提。

[2] 图附录 C-11 表示的博弈场景还有一个平衡点，即行的混合策略 80:20 和列的 40:60。在这种情况下，双方的收益分别均为 3.6，分别取决于矩阵的左上方和右下方。

用，实际上也无法促使双方实现合作；而且，如果所有的符号都具有类似的作用，我们无疑将陷入混淆的迷雾之中。只有在不同特点的暗示符号能够明确指向某个共同选择时，拒绝不同点才意味着失去找到暗示符号或在双方倾向冲突的情况下实现有效结果的可能性。[1]

此外，最符合条件的暗示符号是那些我们排除收益矩阵的数学意义后承认的符号。假设我们在一条垂直交叉的沙漠之路上驾车相向驶往同一个十字路口，没有路标指明正确的方向，彼此之间缺乏好感和信任以及最起码的道德责任感；在这种情况下，一方或许主动减速给对方让路，从而避免在十字路口突然刹车；对方也可能提前预测到这一点。[2]尽管传统的优先机制缺乏法律和道德约束力，但是当合作成为必要时，传统的优先机制仍能发挥积极作用。例如，反对传统优先机制的一方或许会接受其约束力，并意识到他应该感谢这一机制，尽管知道这一机制对其不利以及必须遵守对方对这一标志和预期的认同。正如在第 3 章中我们所分析的，通过这样的逻辑推理，当博弈场景以某个收益矩阵出现或者有利的战略选择被标识为“头”或“尾巴”、 i 、ii、I 、 II 等，图附录 C-10 表示的博弈或许能够获解。

第三者的影响

巧合的是，这些要求选手必须合作的博弈场景——无论是选手利益冲突，还是利益互存——都必须受到调解人的控制和或影响。假设我们授予第三者向原先两名选手传递信息的权力，那么，这个第三者的出现对此前的两名选手将非常有利；甚至当第三者的利益与其他两位选手的战略选择存在利害关系时，他的立场对自己也十分有利。一位慈善的调解人非常有助于完全共同利益博弈的完成，或者是在类似图附录 C-10 的博弈场景中拥有仲裁权力或作为图附录 C-10 博弈的“第三方选手”

[1] “心理优势”反映了博弈选手相互认知的暗示符号的作用。前面案例中的卢斯和雷夫就是利用这一方式处理某种博弈场景中的平衡点问题。详见《博弈论与决策》，第 109~110 页；或者本书之前的脚注对类似博弈场景的论述。

[2] 第 2 章中涉及了有关的利益冲突问题，如两辆汽车在一条狭窄的道路上相向而行。结果表明了这一普遍规律。但是、由于篇幅所限、第 3 章中没有进行详细论述。

的调解人。[1] 在图附录 C-10 中，括号内的值为调解人（信息垄断者）的收益，他的意见对其他两个选手仅仅具有建议性而非强制性。

	Ⅰ	Ⅱ	Ⅲ	Ⅳ
i	5 (2) 6	0	0	0
ii	0	7 (3) 9	0	0
iii	0	0	6 (4) 5	0
iv	0	0	0	9 (1) 7

图附录 C-12

释义收益

最后我们必须指出，这里我们是否将博弈结果看做是客观的可衡量标准（货币、同类物品或是博弈论意义上的“效用”）对推理思维过程并没有影响。只要名义上的收益显而易见，选手能否了解对方的真实意图以及了解的程度如何并不重要。（假设选手知道客观价值或效用价值，并与双方的选择并非成正比，那么“暗示符号”就失去了其应有的作用。混淆或歧义问题将变得更加严重。）

选手数量

本书讨论的博弈场景大多为二人博弈，除了少数情况下涉及具有明确身份的第三者。如果最终的结果取决于选手的共同选择、少数服从多数、多重选择还是成功的联合选择（如第 3 章中涉及的问卷程序），我们讨论的问题将不受到博弈选手数量的限制。如果出现这种情况，歧义问题将变得更加严重，博弈的合作问题更取决于“解决方法”的规律。

[1] 请参阅本书第 3 章，该章涉及如何处理不当得利和自愿调解人的问题。

或许在多人博弈中，合作理论的关系最大，并涉及联合方式问题。因此，对暗示符号的研究和以联合形式出现的沟通渠道将成为联系博弈论和社会学的强大纽带。

结　论

综上所述，合作博弈理论表明，尽管我们就符号属性对某些博弈场景理解得较为透彻，但是我们还无法完全了解默式非零和博弈“严格意义上“的解决方法。鉴于博弈的完全数学模型也可能存在多种暗示符号源，“严格意义上的解决方法”的某些属性也能够成为“理性解决方案”的决定因素之一。这在一定程度上只是一个经验问题，并非仅仅是暗示符号推崇的先验推理问题。

英中文译名对照及索引

A

Accidental war 意外战争 188ff,246ff

Acheson, Dean 迪安·艾奇逊 138n

Agenda 议程 31,51f

Agent 代理人 24,29,33,37,121,142ff

Alchian A.A. 阿尔钦 81n

Allen, Layman E. 莱曼·艾伦 287n

Amster, Warren 沃伦·阿姆斯特 7n

Araitration 仲裁 145,284n

Arms control 军控 5,12,131,136,169n,176,228,230

Atomic weapons 原子武器 见核武器

Auction 拍卖 33,165n

B

Bales R.F. 贝尔斯 86n

Banfield, Edward 班菲尔德·爱德华 127n

Bank robber 银行抢劫犯 139f，见盗窃犯、持枪歹徒、警察

Bargaining 讨价还价 5,21ff,53ff,89n,104,106,227ff,249,262，见博弈、默式谈判

Bargaining power 讨价还价实力 22f

Bavelas, Alex 亚力克斯·贝弗拉斯 86n,172n

Bear by the tail，困难之事 119,149

Beavan, John 约翰·比范 160n

Berlin 柏林 198

Berkowitz M.I. 伯科韦兹 98,93n

Bernard, Jessie 杰西·伯纳德 10n

Bet 赌注 24f

Bilateral monopoly 双边垄断 25 29n35

Binary choice 二元选择 93n

Bishop Robert 罗伯特大主教 287n

Blackmail 勒索，13,44,119,158,159n.见 Extortion
Bluffing 诱骗 23,33,36
Borgatta E.F. 博盖塔 86n
Braithwaite R.B. 布雷恩韦特 113,272n,284n
Brinkmanship 边缘政策 16,199ff
Brodie,Bernard 伯纳德·布罗迪 9n,20,232,239
Burglar 盗窃犯 182,183,207,218 219,229 见持枪歹徒、警察
Burns,Arthur Lee 伯恩斯·亚瑟·李 229n

C

Canyon, Stevenson B. 史蒂文森·坎农 139
Car 汽车 见行路权
Carmichael,L. 卡迈科尔 86n
Cartwright D 卡特赖特 86n,172n
Casuistry 决疑论,诡辩 34,168
Cheating 欺骗 43,45,49,116,135 218n
Chess 国际象棋 58,85,105,106,168
Children 孩子 6,11,13,17,18,136,146,149,170,200
Churchill,Winston S. 邱吉尔 43
Cigarettes 香烟 112,127
Clausewitz 克劳斯韦兹 9
Coalitions 联盟 171f
Cocconi,Guiseppe 吉斯普·科科尼 93n
Collective decision 集体决策 15,220
Collier,John 约翰·科利尔 113n
Collision 碰撞 见行路权
Commitment 承诺 14,24ff,36,43,47,50,121ff,123f,179,184ff,191,195,202,276ff
Communication 沟通 13,14,16,18,26,29,38f,51,53,59,78,84,100,105,115,119,122,124,144,146,149,166,172n,267n,274,277,295,301f
Communist Chinese assets 红色中国的资产 159
Compellent threats 强制威胁 195ff
Compensation 补偿 31,88
Comparative advantage 比较优势 117
Coordination 协作 67,89,92ff,135,172n,283f,289,294ff，见默式谈判
Cross-my-heart 真话 24ff,39n

D

Dare 挑战 112,113,170
Daylight saving 日光节约时间 144
De Gaulle Charles 戴高乐 42,91n
Delegation 代表团 24,29,33,37,121,141ff
Deterrence 威慑 5,6ff,12n,13,119,131,138,187ff,189,195,199,233
Detroit riot 底特律暴动 147
Deutsch,Morton 莫顿·杜切 169n
Diamonds 钻石 94
Disarmament 裁军 见军控

Duffus,R.L. 达弗斯 137n
Dulles,John Foster 杜勒斯 138,159
Duopolists 双边垄断者 31,44

E

Ellsberg,Danie 丹尼尔·埃尔斯博格 113,150n
Elphinstone,Mountstuart 蒙特斯托特·埃尔芬斯通 127n
Empirical study 实证研究 98,113,285
Enforcement 实施 24,44,119ff,122,131ff,134,145
Equilibrium point 平衡点 见博弈论
Esprit de corps 团队精神 92
Etiquette 礼节 91,116n,128,149n
Euthanasia 安乐死 160n
Evidence 根据 15,39,117,124n147ff
Expectation 预期，预期判断 3,101,112,114f,261,279ff,289
Experimental games 实验博弈 见博弈
External economies 外在节约因素 226,229
Extortion 敲诈,勒索 5,13,14,18,31 45n,51,131,141,146 ,见勒索、黑邮件

F

Fads 时尚 91
False alarm, 假警报 188,201,203,220ff,244,246ff,And 见军控、突袭
Fellner,Willam J. 威廉姆·费尔勒 3
Flood,M.M. 弗勒德 164n
Focal points 聚焦点 57,59,68,70,80,111ff,144
Foobail 福特贝尔 110n
Foreign aid 外部援助 11,19n,28,41n
Formaosa 台湾 6,76,115,142,192,265f

G

Games， 博弈 46,56ff,60ff,119ff谈判46ff,99ff,67ff（见谈判）洽作,84沟通见沟通；协调见协调；实验主义的，54ff,109ff,113,162ff,164n,172n 搏弈战术 128n；多重动机 89,102n,103,158,160；非合作式 99,209,291,;正常状态,99,272n 室内 102n ;规则,见转移，强制执行，司法体系；默认，见协调，默式谈判；理论,3n,9f,10n,14,83ff,88,95ff,162ff,267ff,291ff；平衡点,97,210,226,286n,291ff ；数理论，见数理论；最小最大，48,84,96,105,175,271,286n; 弱函数,161；复合战略，见随机化；转移，见转移；随机战略，见随机化；解决方案,154ff 广义解决方案，155; 狭义解决方案，210,275n,291ff 战略,见战略 对称性,见对称性；效用说， 88,115ff,125,287n,302(见价值体系）；零和，见零和博弈。
Gang War 火并 12
Gas warfare 毒气战 75
Gestalt psychology 完形心理学 104,

107, 108n, 164n
Goffman, Erving 欧文·戈夫曼 116n, 128, 149
Gold 黄金 94
Greene, Joseph I 约瑟夫一世 9n
Grodzins, Morton 莫顿·格罗兹斯, 91n
Gromyko 格罗米可 248
Groves, Leslie R. 莱斯利·格罗夫斯 231
Gunman 持枪歹徒 120, 124, 232, see

H

Harassment 折磨迫害 194, 200
Hare, P.H. 黑尔 86n
Harsanyi, John 约翰·哈桑尼 113, 267n, 268, 272n, 278, 279, 280, 284n, 287n
Heise, G.a. 海斯 86n
Hints 暗示 85, 163f, 165n, 295
Hitchcock, Alfred 埃尔弗雷德·希契科克 133n 阿尔福雷德.希区柯克
Hitchhiker 搭便车旅行者 183
Hiltler 希特勒 17
Hogan, H.P. 霍根 86n
Holmes and Moriaty 福尔摩斯，莫日娅蒂 87, 97, 163
Hostages 人质 20, 121, 135, 239, 240n
Hungary 匈牙利 197f

I

Identification 身份 139ff, 145
Immunity grant 豁免 134n
Inadvertent war 意外战争 188ff, 246ff
Incrementalism 渐进主义 170f
India 印度 127n
Indochina 印度支那 76, 259
Institute for Defense Analysis 国防分析研究所 9
Insurance 保险 29, 38, 128, 142
Interdependent decision 互动决策 16, 83ff, 89n

J

Japanese Peace Treaty 《旧金山对日和约》 159

K

Kahn, Herman 赫尔曼·凯恩 120n
Kaplan, Morton A. 莫顿·A. 开普兰 125, 194n
Kaysen, Car l 卡尔·凯森 86n
Keynes, J.M. 凯恩斯 94n
Khrushchev, Nikita, 赫鲁晓夫 17
Kidnapper 绑匪 43, 117, 140
Kissinger, Henry A. 基辛格 171n, 204
Koffka, K. 科夫卡 108n, 110n
Korea 朝鲜 75f

L

Labeling 标识 95, 98, 285, 294ff, 301

Last clear chance 最后良机 37,42,137ff
Law 法律 见司法体系
Leadership 领导 91
Lear,John 约翰·李尔 93n
Leavitt,H.J. 莱维特 86n
Lebanon 黎巴嫩 199
Legal system 司法体系 10,14,18,125n142,149n
Lie detector 测谎仪 109
Limited war 有限战争 5,6,12,53ff,89n,166,167,171,190ff,203,247f,253,257ff
Lippman,Walter 沃尔特·李普曼 38
Little Rock 小石城 146
Logrolling 互投赞成票 33
Lot´ s wife 劳特夫人 110
Loyalty oath 忠诚宣誓 112
Luce,R.Duncan 邓肯·卢斯 87,95n,96n,123,125,181n,210,214,267n,272n,284n,286n,287n,291n,292

M

Machiavelli 马基雅维利 20
Majority 多数派 90,93,97n
Mann,Erwin 欧文·曼 120n
Map 地图 见博弈 ,54f,58f,64,71,99ff,170 财富 135
Marschak,Jacob, 雅戈布·马斯查克 85n,98n
Mathematics 数学 58.65.96.99.105.107,113ff,117,162,285,290,300
Mediator 调停者 63,65,68,71,73,78,143.148.168.171 ,301
Miller G.A 米勒 86n
Minimax 最小最大 见博弈论
Minorant games 弱函数博弈 261
Mixed-motive game 混合动机博弈见博弈
Mixed Strategy 混合战略 见随机化
Mobs 暴民 90,131,146f,71
Moore,O.K. 摩尔 85,93n
Moors, 摩尔人 147
Morgan J.N. 摩根 23
Morgensten,Oskar 奥斯卡·摩根斯顿 87n,116,122n,160,161
Morrison,Philip 菲利普·莫里森 93n
Mossadeq 莫萨德 1318
Moves 行为 24,101,102,117,119ff 160,267ff
Mueller,R.A.H 穆勒 86n
Mutiny 叛乱 90,131,146f,171

N

Nash,John F. 纳什 95n,113,125,267n,271n,272n,283n,284,287,288,289,292n
NATO 北大西洋组织 194
Negroes 非洲黑种人 91
Neighbors 邻居 53,104,116
Noncooperative game 非合作博弈见博弈

Normal form 标准形式 99,272
North Pole 北极 169n,262
Nuclear sharing 核共享 6
Nuclear war 核战争 138
Nuclear weapons 核武器 76,107n,139,171,194,197,257ff

O

Oligopolists 市场寡头 31,44
Open skies 开放天空 119.230.242ff

P

Parachutists 空降兵 54ff,58ff,79. 见图例
Parametric behavior,parametrization， 参变行为,参数化 87 你 4,223ff
Parlor games 室内博弈 102n
Partner,partnership 伙伴，伙伴关系 5,15,89,134,208,215n
Passive resistance 消极抵抗 139
Patterns 格局 67,85,93n,104,164n,168
Penalty 惩罚 128ff,137,150
Pirates 侵权 19
Plate glass window 厚玻璃窗 119,128
Pneumatique 气递信件 148
Police 警察 38,140,146,149,见盗窃犯,诈骗犯,通行权
Positive evidence 有利证据 250
Precedent 先例 34.67 见惯例
Price leadership 价格领先 23
Price stability 价格稳定 74
Price war 价格战 131
Prisoner's dilemma 囚徒困境 98n,213n,214,226
Promise 前提、假设 7,14,43ff,48,50,104,120,129,131ff,141,158n,175,177,186
Pure-collaboration game 完全合作博弈 84、见协作

Q

Quemony 金门 192,265f.见台湾
Quota systems 配额制度 91

R

Racketeer 诈骗犯 12,137n，见勒索，盗窃犯,讹诈,绑架犯,警察
Radner,Roy 罗伊·拉德诺 86n
Raiffa,Howard 霍华德·雷弗 87,95n,96n,123,125,181n,210,214,267n,272n,284n,286n,287n,291,292
Railroad tracks 铁轨 139
RAND Corporation 兰德公司 9
Random strategy 随机战略 见随机化
Randomization 随机化 16,84,88,93,96n,106,130,175ff,187ff,212,219
Rapoport,Anatol 阿纳托尔·拉波波特 7n
Rational behavior 理性行为 3ff,13f,

15,16ff,114,130f,143,144,149,278ff,
291
Rationality 理性 见理性行为
Reaction function 反作用函数 134
151n,158,229
Reprisal 报复行为 见报复
Reputation 声誉 14,25,29,30,36,37
40,127n
Retaliation 报复 12,35,76f,78,107
119,169,194,233ff,252ff,259
Revolt 叛乱 74,90,131,146f,171
Rheingold,Miss 瑞恩高德小姐 94
Richardson,Lewis F 刘易斯·理查森
7n
Riddle 谜语 295
Right of way 行路权 12,21,29,33,35
36n,85,127,144,149,196,300
Riots 暴徒 90,131,146f,171
Roulette 轮盘游戏 197

S

Saar 萨尔河 33
Scenario 意境、场景 169
SchellingT.C. 托马斯·谢林 19n,37n
Secrecy 无记名 29,83,175,176n,见
无记名投票
Secret Ballot 无记名投票 19,38,45n,
148
Shakespeare,William 威廉姆·莎士比亚
140
Sherif,Muzafer 穆扎夫·谢里夫 168
Sherwin.C.W. 舍温 7n,129n
Side payments 单方面支付 31,88
Sikhs 印度锡克教徒 147
Simon,Herbert 赫伯特·西蒙 93
Smithies,Arthur 亚瑟·史密希斯 37
Snyder,Glenn H. 格伦·斯奈德 191n
Solution 解，解决方案 见博弈
Space,communication 空间，沟通 93n
Spies 间谍 20134.148
Status quo 现状 33,64,68,169
Strategy 战略、策略 3,13,15,86,122,
150ff,169n
Strategic move 战略行动 160,见行为
Strike 罢工 5,6,27,35,67，见工会罢
工基金 29
Strike funds 罢工基金
Surprise attack 突袭
Sutherland,E.H 萨瑟兰 141
Symmetry 对称性 114,117,122,131,
267ff
Szilard,Leo 利奥·希拉德 148

T

Tacit bargaining 默式谈判 5,21,54ff,
86,99,156,225ff,261ff,271ff,280n,
288,291 见协作
Tarriff 关税 28,31
Taxes 税收 62,63,65,69,145,170n
Taxis 趋向性 21.119 见行路权
Telephone 电话 85,94,96n,146,147
Test vote 民意表决 112

Threat 威胁 6ff,12,15,35ff,48,50f,104,120,123ff,148,175,177ff,187,

Tipping 颠覆 91n

Tradition 传统 91,92,104,106,107n,260,263,266 见先例

Traffic 交通见警察，通行权

Transfer cost 转移成本 25

Treaties 条约 36,40,142,159 ,见台湾

Trip wire 导火索 6,119,136n,138,187,192,203

Truman,Harry 杜鲁门 43

Trust 信任 14,134f

U

Union 工会 27,29,30,44 ,见罢工

UNRRA 联合国善后救济总署 67

Utilities 效用 88,115ff,125,287n 302,见价值体系

V

Value system 价值观 4,11n,13,16,115ff,171.见效用

Von Neumann,John 约翰·诺埃曼 86n,116,122n,160,161,175

W

Wagner,Harvey 哈维·瓦格纳 272n

Walter,A.A. 沃尔特 86n

War 战争 4ff;威慑见威慑；博弈见图例 绑匪 12；非恶意性,188ff,246ff ;朝鲜 75f；有限性，见有限战争；核，见核武器；价格,131 ；报复，见报复；突袭，见突袭。

Warning system 预警系统 123,220ff,242ff 军控,见军控 突袭见突袭.

Watchdog 监督员 183f

Watchmen 巡逻人员 39n 229

Wohlstetter,Albert J 艾伯特·沃尔斯德特 234n

Zander,A.F. 詹德 86n,172n

Zero-sum game 零和博弈 83ff,87,88ff,96,102n,105,115,119,160f,163,165,175,见博弈论

译后记

《冲突的战略》一书是为了满足理论需要，而非实际外交决策。主要目的是为了建立一个跨学科领域——该领域曾被人们称之为“讨价还价理论”、“冲突理论”或“战略理论”。作者想向人们证明，一些横跨经济学、社会学和政治学，甚至是法学和哲学或者是人类学的基础理论，不仅对从事理论研究的学者大有裨益，而且也有助于人们解决实际问题。与此同时，作者最初认为博弈论（the theory of games）在某些领域应该具有更多实用性。但是现在看来，这也许是一种曲解，因为除了大家比较熟悉的霍华德·雷夫（Howard Raiffa），马丁·舒里克（Martin Shurik）和纳杰尔·霍华德（Nigal Howard）等学者外，博弈理论家更多倾向于将博弈论限制于数学范畴。作者认为，尽管他致力研究的领域可能不会突飞猛进，甚至还尚未有自己的正式名称，但是作者相信将会有更多的有志之士投入到这一领域的完善和发展之中。

感谢在整个翻译过程中对我们的翻译提出过建议、提供了帮助以及督促我们完成进度的朋友、同学和老师。在此感谢赵华、陈之荣、郑明、朱玲、田敏、于洁、陈刚、杨静、李小君、胡倩、谢晓霞、曾琼、曹晓爱、李慧、王俊、赵浩、谢圣哲和谢伦裕对本书翻译工作的帮助，由于译者水平有限，不当之处在所难免，望广大读者批评指正。谢谢！